RATAN TATA LIBRARY,
DELHI UNIVERSITY LIBRARY SYSTEM
WARNING :

Please point out to the Counter Staff physical damage to the book before borrowing, if any

RATAN TATA LIBRARY, (DULS)

(Delhi School of Economics)

Cl. No. B3 2 G9

Ac. No. 68085

This book should be returned on or before the date last stamped below. An overdue charge of 10 Paise per day on general books and 25 Paise per day on text book will be charged for the first two days and 50 Paise from the third day the book is kept overtime.

24 FEB 1995			

ADVANCED CALCULUS

ADVANCED CALCULUS

BY

IVAN S. SOKOLNIKOFF
Professor of Mathematics
University of Wisconsin

INTERNATIONAL STUDENT EDITION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
KŌGAKUSHA COMPANY, LTD.
TOKYO

INTERNATIONAL STUDENT EDITION

Exclusive rights by Kōgakusha Co., Ltd., for manufacture and export from Japan. This book cannot be re-exported from the country to which it is consigned by Kōgakusha Co., Ltd. or by McGraw-Hill Book Company, Inc. or any of its subsidiaries.

I

COPYRIGHT, 1939, BY THE McGRAW-HILL BOOK COMPANY, INC.

All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

PREFACE

There is a pronounced need of a book on advanced calculus that does not sacrifice rigor to such an extent as to become ineffectual as an instrument for developing a critical attitude toward analytical processes, and yet which is sufficiently concrete to be useful to a student with one year of preparation in the calculus. I am under no delusion that this volume completely fills this need, and I shall feel generously repaid for my efforts if it should prove of some aid to those who are faced with the perplexing problem of instruction in analysis.

In preparing this book I have made every effort to keep in mind the difficulties of the reader who is encountering for the first time a serious body of mathematical doctrine. Some ideas that are innately difficult, but whose basic sources stem from geometry, are presented first from an intuitive point of view, so that the essentials can be grasped at once. I did not think it wise to include rigorous arithmetical proofs of such theorems as those on convergence of bounded monotone sequences (Sec. 6), the theorem of Bolzano-Weierstrass (Sec. 7), the theorem of Darboux (Sec. 35), and a few others. This is in accordance with the precept that the most effective means of thwarting interest in mathematics is by misdirecting rigor. A reader who is sufficiently sophisticated to feel the need of arithmetical proofs of these theorems will find them in the treatises to which I refer in the text. The material contained in this volume is so arranged as to minimize the need of irksome references to matters to be established later on. No difficult and essential proofs have been relegated to exercises to be worked out at the reader's leisure.

The subject of advanced calculus is not an easy one, and the working of the problems is essential to a mastery. There are numerous illustrative exercises and problems scattered throughout the text to aid the reader in gaining an insight into the beauty and the wide range of applications of analysis. A student with a good background in the calculus will be able to read this book without omissions.

A considerable portion of the material included here represents a transcription of the lecture notes that I have used from time to time in my classes. These notes place me under some obligation to many of the existing books on analysis, especially to the masterpieces of Goursat, Knopp, and de la Vallée Poussin. I am also much indebted to my wife, who not only undertook the arduous task of reading the manuscript in its various stages, but who also aided me immeasurably with the correction of the proof. Her assistance has resulted in many essential improvements.

IVAN S. SOKOLNIKOFF.

MADISON, WIS.,
November, 1939.

CONTENTS

PAGE
PREFACE v

CHAPTER I

LIMITS AND CONTINUITY

SECTION
1. Number System 1
2. Sequences 3
3. Limit of a Sequence. Convergence 6
4. Inequality of Bernoulli. $\lim_{n \to \infty} \sqrt[n]{a} = 1$. 8
5. Existence of the Limit 10
6. Criterion for Convergence of Monotone Sequences 14
7. Divergent Sequences. Upper and Lower Limits 16
8. Functions of a Single Variable 22
9. Theorems on Limits 26
10. The Base of the Natural Logarithms 28
11. Continuity 31
12. Properties of Continuous Functions 35
13. Uniform Continuity 37

CHAPTER II

DERIVATIVES AND DIFFERENTIALS

14. Derivatives 41
15. Differentials 43
16. Derivatives of Composite Functions 45
17. Derivatives and Differentials of Higher Orders 47
18. Fermat's Theorem 49
19. Rolle's Theorem 50
20. Mean-value Theorem 51
21. Theorem of Cauchy. L'Hospital's Rule 53

CHAPTER III

FUNCTIONS OF SEVERAL VARIABLES

22. Limits and Continuity 58
23. Partial Derivatives 62
24. Differentiation of Composite Functions 67
25. Differentiation of Composite and Implicit Functions 71
26. Euler's Theorem 75
27. Directional Derivatives 76
28. Tangent Plane and Normal Line to a Surface 80
29. Space Curves 83
30. Directional Derivatives in Space 86

SECTION PAGE
31. Partial Derivatives of Higher Order 87
32. Higher Derivatives of Implicit Functions 89
33. Change of Variables 91

CHAPTER IV

DEFINITE INTEGRALS

34. Riemann Integral 99
35. Riemann Integral—(*Continued*) 104
36. Direct Evaluation of Integrals 110
37. Mean-value Theorems for Integrals 113
38. Fundamental Theorem of the Integral Calculus 118
39. Differentiation under the Integral Sign 121
40. Change of Variable 124
41. Applications of Definite Integrals 126

CHAPTER V

MULTIPLE INTEGRALS

42. Double Integrals 130
43. Evaluation of the Double Integral 131
44. Geometric Interpretation of the Double Integral 139
45. Triple Integrals 142
46. Change of Variables in a Double Integral 147
47. Transformation of Points 153
48. Change of Variables in a Triple Integral 155
49. Spherical and Cylindrical Coordinates 157
50. Surface Integrals 161
51. Green's Theorem in Space 167
52. Symmetrical Form of Green's Theorem 170

CHAPTER VI

LINE INTEGRALS

53. Definition of Line Integral 174
54. Area of a Closed Curve 178
55. Green's Theorem for the Plane 181
56. Properties of Line Integrals 185
57. Multiply Connected Regions 192
58. Line Integrals in Space 195
59. Stokes's Theorem 196
60. Applications of Line Integrals 199

CHAPTER VII

INFINITE SERIES

61. Infinite Series 209
62. Series of Positive Terms 213
63. More General Tests 225
64. Series of Arbitrary Terms 233
65. Absolute Convergence 236
66. Properties of Absolutely Convergent Series 240

Section Page
67. Double Series 246
68. Series of Functions. Uniform Convergence 247
69. Geometric Interpretation of Uniform Convergence 252
70. Properties of Uniformly Convergent Series. 256
71. Weierstrass Test for Uniform Convergence. 262
72. Abel's Test for Uniform Convergence. 264

CHAPTER VIII

POWER SERIES

73. Power Series. 267
74. Interval of Convergence. 269
75. Properties of Functions Defined by Power Series. 275
76. Abel's Theorem 276
77. Uniqueness Theorem on Power Series. 279
78. Algebra of Power Series. 280
79. Calculations Involving Power Series 285

CHAPTER IX

APPLICATIONS OF POWER SERIES

80. Extended Law of the Mean 291
81. Taylor's Formula. 293
82. Taylor's Series. 296
83. Applications of Taylor's Formula. 298
84. Euler's Formulas and Hyperbolic Functions. 306
85. Integration of Power Series 309
86. Evaluation of Definite Integrals 310
87. Maxima and Minima of Functions of One Variable. 315
88. Taylor's Formula for Functions of Several Variables 317
89. Maxima and Minima of Functions of Several Variables. 321
90. Constrained Maxima and Minima 327
91. Lagrange's Multipliers 331

CHAPTER X

IMPROPER INTEGRALS

92. Integral with Infinite Limit 335
93. Tests for Convergence of Integrals with Infinite Limits 341
94. Integrals in Which the Integrand Becomes Infinite. 347
95. Tests for Convergence of Integrals Whose Integrands Become Infinite 350
96. Operations with Improper Integrals. 352
97. Evaluation of Improper Integrals. 357
98. Improper Multiple Integrals. 365
99. Gamma Functions 372

CHAPTER XI

FOURIER SERIES

100. Criterion of Approximation 378
101. Fourier Coefficients. 380

SECTION PAGE
102. Conditions of Dirichlet 385
103. Orthogonal Functions. 388
104. Expansion of Functions in Fourier Series 390
105. Sine and Cosine Series 397
106. Extension of Interval of Expansion. 401
107. Complex Form of Fourier Series 403
108. Differentiation and Integration of Fourier Series 405
109. Fourier Integral 409

CHAPTER XII

IMPLICIT FUNCTIONS

110. A Simple Problem in Implicit Functions. 415
111. Generalization of the Simple Problem. 418
112. Functional Dependence. 423
113. Existence Theorem for Implicit Functions. 425
114. Existence Theorem for Simultaneous Equations 430
115. Functional Dependence. 433
116. Properties of Jacobians 438

INDEX. 441

ADVANCED CALCULUS

CHAPTER I

LIMITS AND CONTINUITY

1. Number System. This section consists of a brief summary of the concepts and properties of the system of numbers that will be considered in this volume. The starting point is the system of positive integers. This system permits the operations of addition and multiplication but proves inadequate to permit the solution of many simple linear equations. Thus, if a, b, c, and d are positive integers, $a + x = b$ and $cy = d$ do not possess solutions that are positive integers unless $b > a$ and d is an integral multiple of c. The first equation serves as the definition of the new number zero if $b = a$ and of the negative integers if $b < a$. The second equation demands the introduction of the fractional numbers $\frac{d}{c}$.

The system of numbers consisting of the positive and negative integers and fractions and zero is called the *system of rational numbers.* It admits the four fundamental operations of addition, subtraction, multiplication, and division (except by zero). Moreover, it is "ordered," which means that if a and b are two different numbers, then $a > b$ or $b > a$. Also, if c is a third rational number and $a > b$ and $b > c$, then $a > c$. The system of rational numbers possesses the additional property of density, that is, between any two different rational numbers there is always another rational number (and, therefore, an infinite number of rational numbers) that is greater than one of the numbers and less than the other.

A further extension of the number system is required when it becomes necessary to solve equations of the second and higher degrees. In elementary algebra a rational number is defined as one that can be expressed in the form $\frac{a}{b}$, $(b \neq 0)$, where a and b

are integers, and an irrational number is defined as one that does not possess this property. The irrational numbers include not only such numbers as $\sqrt{2}$, $\sqrt[3]{4}$, $\sqrt[4]{3}$, etc., which arise in the solution of algebraic equations, but also π, e, etc. The method of introducing the irrational numbers, which is outlined below, is due to Dedekind.*

Assume that all the rational numbers have been separated into two sets in such a manner that every number in the first set is less than every number in the second set. A separation of this type is called a *cut*, or *partition*, (French, *coupure;* German, *Schnitt*) of the number system. Four cases present themselves, but it is easy to see that one of these is impossible. They are:

i. The first set, A, contains a greatest element a and the second set, B, contains a smallest element b. It cannot be that $a = b$ since every element of A is less than every element of B. But if $a < b$, then there are rational numbers which are greater than a and less than b (by the property of density). These numbers, being greater than a, must belong to B, and at the same time, being less than b, must belong to A, which is impossible.

ii. A contains a greatest element a, and B contains no smallest element. For example, let A consist of the number 2 and all rational numbers less than 2, whereas B consists of all rational numbers greater than 2.

iii. B contains a smallest element b, and A contains no greatest element. For example, let B contain 2 and all rational numbers greater than 2, and A contain all rational numbers less than 2.

iv. A contains no greatest element, and B contains no smallest element. For example, let A consist of all the negative rational numbers and all positive rational numbers whose square is less than 2, and let B consist of all the positive rational numbers whose square is greater than 2. Then the cut defines the irrational number $\sqrt{2}$. In general, let the system of rational numbers be separated into two sets A and B in such a way that *every number in A is less than every number in B and A contains no greatest element and B contains no smallest element.* Then the cut defines an irrational number that has the property of being greater than all the elements in A and less than all the elements in B.

The system of numbers that includes all the rational numbers and all the irrational numbers is called the *real number system.*

---

* Stetigkeit und irrationale Zahlen, Braunschweig, 1872.

The fundamental operations of addition and multiplication and the inverse operations of subtraction and division are defined for irrational numbers in such a way* that the system of real numbers possesses the commutative, associative, and distributive properties. Moreover, if the product of any two real numbers is zero, then at least one of the numbers must be zero.

It is often desirable to have a correspondence between the real numbers and the points of a straight line, so that analytic methods can be applied to geometry. In order to effect a one-to-one correspondence, it is necessary to assume the *axiom of Cantor-Dedekind*, which states that "to each point on the line there corresponds one and only one real number and, conversely, to each real number there corresponds one and only one point on the line."

As is already familiar from elementary algebra, the solution of the general quadratic equation $ax^2 + bx + c = 0$, in the case where the discriminant $b^2 - 4ac$ is negative, necessitates the introduction of a new type of number of the form $u + iv$, where u and v are real numbers and i is a number such that $i^2 = -1$. Such numbers are called *complex numbers*. The system of complex numbers includes the system of real numbers (as a special case when $v = 0$), and the fundamental operations must be defined so that they are consistent with those already in use for real numbers. This can be done by considering the complex numbers in the form $u + iv$ or by treating them as ordered pairs (u, v) of real numbers. Since this volume deals primarily with real numbers, these definitions and the properties of complex numbers will not be discussed here.†

2. Sequences. Let some process of construction yield a succession of real numbers

$$x_1, x_2, x_3, \cdots, x_n, \cdots,$$

where it is assumed that every term x_i is followed by other terms. A succession of numbers formed in this way is called a *sequence* and is denoted by the symbol $\{x_i\}$. The individual terms of the sequence $\{x_i\}$ are called the *elements of the sequence.*

---

* De la Vallée Poussin, C. J., Cours d'analyse infinitésimale, vol. 1, pp. 1–8; Fine, H. B., College Algebra, pp. 39–55; Hobson, E. W., Theory of Functions of a Real Variable, vol. 1, pp. 20–34.

† See De la Vallée Poussin, C. J., Cours d'analyse infinitésimale, vol. 1, pp. 37–42; Fine, H. B., College Algebra, pp. 70–78; Burkhardt, H., and Rasor, S. E., Theory of Functions of a Complex Variable, pp. 1–23.

Examples of sequences are

$$\begin{aligned}
&(a)\ 1, 2, 3, 4, \cdots, n, \cdots;\\
&(b)\ 1, -2, 3, -4, \cdots, (-1)^{n-1}n, \cdots;\\
&(c)\ \frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, -\frac{1}{16}, \cdots, (-1)^{n-1}\frac{1}{2^n}, \cdots;\\
&(d)\ 0, 1, 0, 1, \cdots, \tfrac{1}{2}[1 + (-1)^n], \cdots;\\
&(e)\ 2, 3, 5, 7, 11, \cdots, p_n, \cdots.
\end{aligned}$$

It is not essential that the general term of the sequence be given by some simple formula, as is the case in the first four examples above. The sequence (*e*) represents the succession of prime numbers, and p_n stands for the nth prime number. There is no formula available for the determination of the nth prime number, but it is possible to calculate all prime numbers less than any given number N. Although the process of constructing the sequence (*e*) is much more complicated than the processes used for the other examples, nevertheless it yields a definite succession of numbers.

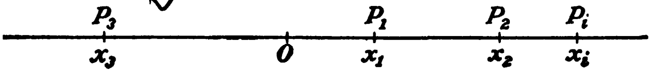

FIG. 1.

It is convenient to represent the elements x_i of the sequence $\{x_i\}$ by points P_i on a straight line. The points P_i are so chosen that the distance from an arbitrary point O of the line to the point P_i is equal to x_i. In this way a correspondence is established between the elements of the sequence $\{x_i\}$ and a set of points on a straight line. To avoid circumlocution, the elements x_i of the sequence will be called *points*, instead of saying that they are numbers such that the distances of the points P_i corresponding to them, from some fiducial point O, are equal to x_i (Fig. 1).

In studying the behavior of sequences of real numbers, or sets of points corresponding to them, several interesting cases present themselves. It may happen that one can find a positive number M such that the inequality

$$|x_n| \leq M$$

or

$$-M \leq x_n \leq M$$

is satisfied for every n, however large. In such a case the sequence $\{x_i\}$ is said to be *bounded.* A geometrical interpretation of this circumstance is that the points P_i corresponding to the elements of the sequence always remain within a segment of length $2M$ with the point O as the midpoint of the segment.

As an example, consider a succession of values

$$\frac{1}{2}, -\frac{2}{3}, \frac{3}{4}, -\frac{4}{5}, \ldots, (-1)^{n-1}\frac{n}{n+1}, \ldots .$$

It is clear that, no matter how large n is taken,

$$|x_n| = \frac{n}{n+1} < 1,$$

so that the sequence $\left\{(-1)^{n-1}\frac{n}{n+1}\right\}$ is bounded.

The sequences (a), (b), and (e) (p. 4) are unbounded, but any number greater than or equal to ½ will serve as a bound for the sequence (c), and any number greater than or equal to unity is a bound for (d).

A sequence $\{x_i\}$ is called a null sequence if subsequent to the choice of a positive number ϵ, however small, one can find a positive integer p such that $|x_n| < \epsilon$ for all values of $n > p$.

If a graphical mode of representing the sequence by points on a straight line be adopted, the definition of the null sequence means that the points P_{p+1}, P_{p+2}, . . . , corresponding to x_{p+1}, x_{p+2}, . . . , will all lie within the interval of width 2ϵ with the point O as midpoint.

It is important to observe that the integer p in the foregoing definition is determined after the choice of the number ϵ has been made, and hence, p depends on the magnitude of ϵ. To clarify this, consider the sequence $\left\{\frac{1}{n^2}\right\}$, or

$$\frac{1}{1^2}, \frac{1}{2^2}, \frac{1}{3^2}, \ldots, \frac{1}{n^2}, \ldots .$$

This sequence is a null sequence, since one can always find an integer p such that for all values of n greater than p

$$\frac{1}{n^2} < \epsilon,$$

or

$$n^2 > \frac{1}{\epsilon}.$$

Hence, if one chooses for p any positive integer greater than the number $\sqrt{\frac{1}{\epsilon}}$, the inequality will be satisfied. If ϵ is chosen as $1/100$, any integer greater than 10 will serve as p; but if $\epsilon = \frac{1}{10{,}000}$, p must be chosen greater than 100.

PROBLEMS

1. Show that $\left\{\frac{1}{n}\right\}$ is a null sequence.

2. Prove that $\left\{\frac{1}{10^n}\right\}$ is a null sequence. Show that n must be chosen greater than $\log_{10}\frac{1}{\epsilon}$ in order that the inequality $|x_n| < \epsilon$ be satisfied.

3. Limit of a Sequence. Convergence. *If $\{x_i\}$ is a given sequence and there exists a number L such that the sequence $\{x_i - L\}$ is a null sequence, then the given sequence is said to be convergent and the elements x_i of the sequence are said to approach the limit L.*

Recalling the definition of the null sequence, it is clear that if the limit of the sequence $\{x_i\}$ is L, then corresponding to any positive number ϵ, a positive integer p can be assigned such that

$$|L - x_n| < \epsilon, \qquad \text{for all} \qquad n > p.$$

The last statement is frequently written in the following ways:

$$\lim_{n\to\infty} x_n = L,$$

or

$$x_n \to L,$$

when $n \to \infty$. The symbol $n \to \infty$ is read "n tends to infinity."

It is clear that a null sequence is a convergent sequence with the value zero as the limit of the sequence. Moreover, it is obvious that every convergent sequence is bounded.*

---

* The converse of this statement is not true. Thus, the sequence 0, 1, 0, 1, 0, . . . is bounded but not convergent.

The following theorem regarding null sequences is established easily:

Theorem. *If the terms of any null sequence*

$$x_1, x_2, x_3, \cdots, x_n, \cdots$$

are multiplied by a sequence of bounded factors

$$a_1, a_2, a_3, \cdots, a_n, \cdots,$$

then the resulting sequence

$$a_1x_1, a_2x_2, a_3x_3, \cdots, a_nx_n, \cdots$$

is a null sequence.

Since the numbers a_n are bounded, there exists a positive number M, independent of n, such that

$$|a_n| < M. \tag{3-1}$$

Moreover, since the sequence $\{x_i\}$ is a null sequence, it is possible to find for every $\epsilon > 0$ a positive integer p such that

$$|x_n| < \frac{\epsilon}{M}, \tag{3-2}$$

whenever $n > p$. Multiplying the inequalities (3-1) and (3-2) gives

$$|a_nx_n| < \frac{\epsilon}{M}M = \epsilon,$$

which establishes the fact that the sequence $\{a_nx_n\}$ is a null sequence.

Inasmuch as the notion of the limit of a sequence is one of the basic concepts of analysis, it is desirable to consider an example.

Illustrative Example. Let the variable x assume a set of values

$$x_1 = 0.3, \qquad x_2 = 0.33, \qquad x_3 = 0.333, \cdots.$$

It will be shown that

$$\lim_{n\to\infty} x_n = \tfrac{1}{3};$$

that is, corresponding to an arbitrary $\epsilon > 0$, one can determine a positive integer p, such that the difference

$$|L - x_n| = |1/3 - x_n|$$

becomes and remains less than ϵ for all values of $n > p$.

Note that

$$\frac{1}{3} - x_1 = \frac{1}{30}, \quad \frac{1}{3} - x_2 = \frac{1}{300}, \ \cdots, \frac{1}{3} - x_n = \frac{1}{3 \cdot 10^n}, \ \cdots$$

Now choose an ϵ, and impose the demand that

$$\frac{1}{3} - x_n = \frac{1}{3 \cdot 10^n} < \epsilon.$$

But the inequality is equivalent to

$$3 \cdot 10^n > \frac{1}{\epsilon},$$

and taking logarithms to the base 10 of both sides of this inequality gives*

$$\log 3 + n > \log \frac{1}{\epsilon},$$

or

$$n > -(\log \epsilon + \log 3) = -\log 3\epsilon.$$

Thus, if p is chosen as any integer greater than $|-\log 3\epsilon|$, the inequality will be satisfied for all values of n greater than this particular value of p.

A geometrical interpretation of the statement

$$\lim_{n \to \infty} x_n = L$$

is that all but a finite number of the points x_i will fall within the interval of width 2ϵ about the point L.

The totality of points whose distance from a given point L is less than a given number ϵ is called a *neighborhood* of this point. The term *vicinity* is also used to denote the neighborhood.

4. Inequality of Bernoulli. $\lim\limits_{n \to \infty} \sqrt[n]{a} = 1$**.** Let it be required to determine the limit of $\sqrt[n]{a}$ as $n \to \infty$ by assuming the sequence of positive integral values and where a is any positive number.

---

[*] It will be noted that if $A > B$, $\log A > \log B$. This fact follows directly from the definition of the logarithm.

This limit will be determined with the aid of an inequality established by James Bernoulli, which states that if x is any real number such that

$$x > -1 \text{ and } x \neq 0,$$

then

$$(1 + x)^n > 1 + nx,$$

where n is any integer greater than unity.

This inequality is obviously true if $n = 2$, since

$$(1 + x)^2 = 1 + 2x + x^2 > 1 + 2x.$$

The proof for any positive integral value of n will be established by the method of mathematical induction. Assume that the inequality is true for $n = k$, and prove that then it is valid for $n = k + 1$. Since the inequality is known to be true for $n = 2$, it will follow that it is true for $n = 3, 4, 5, \cdots$.

Assuming the truth of the inequality for $n = k$, that is,

$$(1 + x)^k > 1 + kx,$$

then

$$\begin{aligned}(1 + x)^k(1 + x) > (1 + kx)(1 + x) &= 1 + (k + 1)x + kx^2\\ &> 1 + (k + 1)x,\end{aligned}$$

and the proof is complete.

To show that $\lim_{n\to\infty} \sqrt[n]{a} = 1$ for $a > 0$, consider the two cases:

i. $a \geq 1$,
ii. $0 < a < 1$.

i. If $a = 1$, the statement is obvious.

If $a > 1$, $\sqrt[n]{a} > 1$, so that one can write it as

$$\sqrt[n]{a} = 1 + x_n. \tag{4-1}$$

Hence,

$$a = (1 + x_n)^n,$$

and by the inequality of Bernoulli

$$a = (1 + x_n)^n > 1 + nx_n > nx_n.$$

Consequently,

$$x_n < \frac{a}{n},$$

so that the numbers x_n form a null sequence, and it is seen from (4-1) that the sequence of numbers

$$\sqrt[n]{a} \to 1 \qquad \text{as} \qquad n \to \infty.$$

ii. If $0 < a < 1$, then $\frac{1}{a} > 1$, and from (i) it follows that

$$\left\{\sqrt[n]{\frac{1}{a}} - 1\right\} \tag{4-2}$$

is a null sequence. If the terms of the sequence (4-2) are multiplied by the factors

$$a, \sqrt{a}, \sqrt[3]{a}, \cdots, \sqrt[n]{a}, \cdots,$$

there results the sequence

$$\{1 - \sqrt[n]{a}\},$$

which, by the theorem of Sec. 3, is a null sequence since the factors $\sqrt[n]{a}$ are bounded.*

PROBLEM

Find the limits of the following sequences:

(a) $0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \cdots, \frac{n-1}{n}, \cdots;$

(b) $\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \cdots, \frac{n+1}{n}, \cdots;$

(c) $1, \frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{2^n}, \cdots;$

(d) $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt[3]{2}}, \frac{1}{\sqrt[4]{2}}, \cdots, \frac{1}{\sqrt[n]{2}}, \cdots;$

(e) $\frac{1}{1}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \cdots, \frac{1}{\sqrt{n}}, \cdots.$

5. Existence of the Limit. The definition of convergence of a sequence $\{x_i\}$ given in Sec. 3 implies the existence of the limit L, since the application of the test, based on this definition, requires the knowledge of the limit L.

---

* The factors $\sqrt[n]{a}$ are bounded since $a \leq \sqrt[n]{a} < 1$.

The problem of the numerical calculation of the limit of any given sequence ordinarily is an exceedingly difficult one, but in a great variety of practical and theoretical investigations one is satisfied with knowledge of the behavior of the sequence as to convergence and does not require the numerical value of the limit. There is a profound theorem due to A. Cauchy that enables one to establish the existence of the limit of a sequence without actually calculating the limit itself. This important theorem is as follows:

The Fundamental Principle of Convergence. *A necessary and sufficient condition for the existence of the limit of the sequence* $\{x_n\}$ *is that for any* $\epsilon > 0$ *one can find a positive integer* N *such that for any pair of indices* m *and* n*, both greater than or equal to* N,

$$|x_m - x_n| < \epsilon.$$

It will be established first that the condition enunciated in the principle is a necessary one. Let the sequence converge, and denote its limit by L. Then for any positive ϵ, however small, one can find a positive integer N such that

$$|L - x_n| < \frac{\epsilon}{2} \tag{5-1}$$

whenever $n \geq N$. But one can write

$$x_m - x_n = (x_m - L) + (L - x_n)$$

and, noting that the absolute value of the sum is less than or at most equal to the sum of the absolute values, one has

$$\begin{aligned} |x_m - x_n| &\leq |x_m - L| + |L - x_n| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

where the last step results from the hypothesis (5-1).

The demonstration of the sufficiency of the condition is not so easy. An analytical proof makes considerable demands on the mathematical equipment of the reader,[*] but the correctness of the converse of the principle can be seen from the following geometrical considerations, which in essence do not depart from a rigorous analytical proof.

---

[*] See HOBSON, E. W., Theorv of Functions of a Real Variable, vol. 1, p. 38.

The hypothesis this time is that for any $\epsilon > 0$,

$$|x_m - x_n| < \epsilon, \tag{5-2}$$

whenever m and n both exceed or are equal to N, which, of course, depends on the magnitude of ϵ. Choose $\epsilon = \frac{1}{2}$, and denote the corresponding value of N by N_1; then if p_1 is any index greater than N_1,

$$|x_n - x_{p_1}| < \tfrac{1}{2}$$

for every $n > p_1$.

Geometrically this means that all points x_n of the sequence for which $n > p_1$ lie within the segment of length unity which has x_{p_1} as its midpoint (Fig. 2).

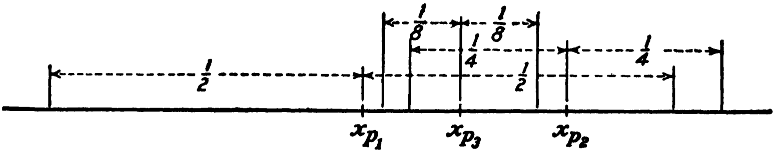

Fig. 2.

If ϵ is taken equal to $\frac{1}{4}$, there will exist an integer $p_2 > p_1$, such that for every $n > p_2$

$$|x_n - x_{p_2}| < \frac{1}{2^2}.$$

Then all points x_n, for $n > p_2$, will lie within that portion of the segment of length $\frac{1}{2}$, with x_{p_2} as the midpoint, which is contained in the segment of length unity. Setting $\epsilon = \frac{1}{8}$, one is assured of the existence of an integer $p_3 > p_2$, such that

$$|x_n - x_{p_3}| < \frac{1}{2^3},$$

whenever $n > p_3$. That is, the inequality is satisfied for all points x_n that are contained in that portion of the interval of length $\frac{1}{4}$, with x_{p_3} as the midpoint, which is common to the interval of length $\frac{1}{2}$ with x_{p_1} as the midpoint.

Continuing this process by assigning to ϵ successively the values $\frac{1}{2^4}$, $\frac{1}{2^5}$, etc., one builds up a sequence of intervals diminishing in length and such that the end points of these intervals tend to a definite point L, which is the limiting point of the sequence. Thus the existence of the limit L is demonstrated, provided the condition (5-2) is satisfied.

The fundamental criterion of convergence is frequently stated in the following somewhat different, but equivalent, form, which the reader will have no difficulty in deducing.

Theorem. *A necessary and sufficient condition for the convergence of the sequence*

$$x_1, x_2, \cdots, x_n, \cdots$$

is that for any $\epsilon > 0$, *one can find a positive integer* N *such that*

$$|x_{n+k} - x_n| < \epsilon$$

when $n \geq N$, *and for every positive integer* k.

It is seen from the foregoing that the significance of the criterion is, roughly, that all terms of the sequence for sufficiently large values of the indices lie arbitrarily near one another.

In order to illustrate the application of the fundamental criterion, consider the sequence

$$1, \tfrac{1}{2}, \tfrac{3}{4}, \tfrac{5}{8}, \ldots,$$

where each term is the arithmetic mean of the two terms which immediately precede it. Thus the general formula for the nth term of the sequence ($n > 2$) is

$$x_n = \frac{x_{n-1} + x_{n-2}}{2}.$$

It is easily seen that the difference between two consecutive terms of the sequence is

$$x_{n+1} - x_n = \frac{(-1)^n}{2^n}.$$

But all the terms following x_{n+1} lie between x_n and x_{n+1}, and if any $\epsilon > 0$ has been assigned, one can choose a positive integer N so great that

$$\frac{1}{2^N} < \epsilon$$

With this choice of N the inequality

$$|x_m - x_n| < \epsilon$$

will surely be satisfied for any m and n that are not less than N. Thus, by the fundamental criterion, the sequence is convergent.

As an exercise the reader may attempt to show by induction that

$$\left|x_n - \frac{2}{3}\right| = \frac{1}{3} \cdot \frac{1}{2^{n-1}},$$

so that the limit of this sequence is $2/3$, although one need not know the limit itself in order to establish the convergence of the sequence.

6. Criterion for Convergence of Monotone Sequences. The fundamental criterion of convergence is applicable to any type of sequence, but if one is dealing with monotone sequences (that is, such that either $x_{n+1} \geq x_n$ for all values of n or $x_{n+1} \leq x_n$ for all values of n), there is a simpler criterion that states that *a monotone sequence which is bounded is convergent.*

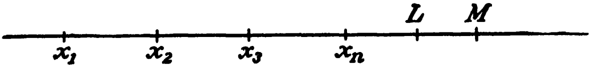

FIG. 3.

In order to make this assertion plausible, assume the sequence $\{x_n\}$ to be monotone increasing and bounded, that is,

$$x_{n+1} \geq x_n,$$

and $|x_n| < M$ for every value of n. Geometrically this means that the points corresponding to the elements x_n of the sequence move to the right with increasing values of the index n (Fig. 3) but are always restricted to lie to the left of the point corresponding to the number M. It is thus intuitively clear that there must be some point $L \leq M$ toward which the points x_n tend with increasing n. This point L is the limit of the sequence.

The discussion in the preceding paragraph applies to the case of monotone decreasing sequences bounded on the left if the words *right* and *left* are interchanged. It should be remarked that the foregoing discussion is based on an intuitive notion as to what happens in a corresponding geometrical situation. A rigorous analytical proof based on a refinement of this pictorial idea is omitted.

To illustrate the use of this criterion of convergence of monotone sequences, consider the sequence $\{x_n\}$, where

$$\begin{aligned} x_1 &= \tfrac{1}{2}, \\ x_2 &= \tfrac{1}{3} + \tfrac{1}{4}, \\ x_3 &= \tfrac{1}{4} + \tfrac{1}{5} + \tfrac{1}{6}, \\ &\cdots\cdots\cdots\cdots \\ x_n &= \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

This sequence is monotone increasing since

$$x_{n+1} - x_n = \frac{1}{2n+1} - \frac{1}{2n+2} > 0.$$

Moreover, it is bounded for

$$x_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \leq n \cdot \frac{1}{n+1} < 1.$$

Therefore, the sequence $\{x_n\}$ is convergent.

The power of this criterion lies in the fact that it requires only the proof of monotonicity and of the boundedness of the sequence. Accordingly, it is much easier to apply than the more general fundamental criterion.

PROBLEMS

1. Show that the sequence

$$1, 1 + \frac{1}{2}, 1 + \frac{1}{2} + \frac{1}{4}, \cdots, 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n}, \cdots$$

is convergent.

2. Prove the convergence of the following sequences:

(a) $\left\{\frac{n-1}{n}\right\}$;

(b) $\left\{\frac{1}{2^n}\right\}$;

(c) $\left\{\frac{n+1}{n}\right\}$;

(d) $\left\{\dfrac{1}{\sqrt[n]{2}}\right\}$;

(e) $\left\{\dfrac{n}{2n+1}\right\}$.

7. Divergent Sequences. Upper and Lower Limits.

Consider again a variable x that assumes a set of real values

$$x_1, x_2, \cdots, x_n, \cdots$$

and a set of points $P_1, P_2, \cdots, P_n, \cdots$ associated with it. If the sequence $\{x_n\}$ fails to converge, it is said to *diverge*.

The sequence may fail to approach a limit because the variable x becomes infinite, which simply means that one can find a positive integer N such that for every $n \geq N$

$$|x_n| > M,$$

where M is a preassigned positive number, however large. In geometrical language, the statement that x becomes infinite means that no matter how large a segment with O as the midpoint one may take, there will be points P_n that lie outside this segment. If the variable x becomes infinite in such a way that after some value, say x_p, it remains positive, it is said to *tend to positive infinity*. On the other hand, if it becomes and remains negative after some x_p, it is said to become *negatively infinite* or to *tend to negative infinity*.

For example, if x takes on the set of values

$$-1, -2, -3, \cdots, -n, \cdots,$$

it becomes negatively infinite. This behavior is commonly denoted by writing

$$\lim_{n\to\infty} x_n = -\infty$$

but this notation is bad since in this case the limit does not exist. It should be kept clearly in mind that the mark ∞ is not a number but a symbol representing the idea that the variable x increases without limit.

The variable x may also change in such a way that it neither approaches a limit nor becomes infinite. For example, if x assumes a sequence of values

$$1, -1, 1, -1, \cdots,$$

it is clearly bounded, but instead of approaching a limit, it oscillates between $+1$ and -1.

In defining the limit L of a sequence $\{x_i\}$, it was stated that the inequality

$$|L - x_n| < \epsilon$$

must be fulfilled for every value of $n > p$, regardless of how small the number ϵ is chosen. A geometrical interpretation of this statement is that an infinite number of points P_i, corresponding to the elements x_i of the sequence, lie in every interval of width 2ϵ whose midpoint P corresponds to the limit L of the sequence. Accordingly, the point P may be called *the point of condensation* or *the limiting point of the sequence* $\{x_i\}$.

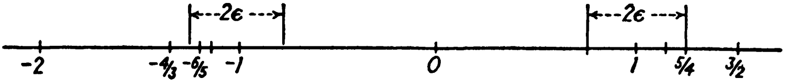

FIG. 4.

Points of condensation may arise even if the sequence $\{x_i\}$ is not convergent. For example, the sequence

$$-2, \frac{3}{2}, -\frac{4}{3}, \frac{5}{4}, -\frac{6}{5}, \cdots, (-1)^n\frac{n+1}{n}, \cdots \tag{7-1}$$

does not converge, but there are infinitely many elements of (7-1) in the vicinity of the points -1 and $+1$. Any interval of width 2ϵ (Fig. 4) about $x = -1$ or $x = +1$ contains infinitely many terms of the given sequence, so that the points -1 and $+1$ are limiting points of the sequence (7-1).

The points of condensation are of considerable importance in the study of the so-called *power series.*[*]

Definition of Limiting Point. *A number L is called a limiting point of a sequence $\{x_i\}$ if for any $\epsilon > 0$ there is an infinite number of terms of the sequence satisfying the inequality*

$$|x_n - L| < \epsilon.$$

The distinction between the definition of the limiting point and that of the limit is that, in the definition just given, the inequality

$$|x_n - L| < \epsilon$$

---

[*] See Chap. VIII.

must be satisfied for *any* infinite number of terms of the sequence, whereas in the definition of the limit the same inequality must be fulfilled for *every* x_n *whose index n exceeds some fixed positive integer p.* Any number that occurs an infinite number of times in a sequence is obviously a limiting point. Thus, the sequence

$$1, \tfrac{1}{2}, 1, \tfrac{1}{3}, 1, \tfrac{1}{4}, \ldots \tag{7-2}$$

has 1 and 0 as its limiting points, but the limit of this sequence obviously does not exist.

A fundamental theorem concerning bounded sequences, due to B. Bolzano and K. Weierstrass, is stated as follows:

Every bounded sequence possesses at least one limiting point. Moreover, it can be established that every bounded sequence has a least limiting point and a greatest limiting point.* This latter assertion is certainly obvious if the number of limiting points of the sequence is finite.

Definitions of Upper and Lower Limits. *The least limiting point of a bounded sequence* $\{x_n\}$ *is called the lower limit and is denoted by the symbol*

$$\underline{\lim}\, x_n = l.$$

The greatest limiting point of a bounded sequence $\{x_n\}$ *is called the upper limit and is designated by the symbol*

$$\overline{\lim}\, x_n = L.$$

Recalling the definition of the limiting point, it is clear that if l is the lower limit of the sequence $\{x_i\}$, then for any $\epsilon > 0$, there will be infinitely many terms of the sequence such that

$$x_n < l + \epsilon,$$

and at most a finite number of the x_n's such that

$$x_n < l - \epsilon.$$

On the other hand, L is the upper limit of the sequence $\{x_n\}$ if there are infinitely many terms of the sequence such that

$$x_n > L - \epsilon,$$

---

* For a rigorous arithmetical proof of these assertions see E. W. Hobson, Theory of Functions of a Real Variable, Secs. 46–48.

and only a finite number (or none at all) such that

$$x_n > L + \epsilon.$$

It follows from these statements that -1 is the lower limit of the sequence (7-1), while $+1$ is its upper limit. The lower and upper limits of the sequence (7-2) are 0 and 1, respectively.

It is customary to extend the definitions of the upper and lower limits to unbounded sequences by saying that if there are infinitely many terms of the sequence $\{x_i\}$ whose magnitude exceeds any preassigned positive number M, then the upper limit L is $+\infty$; on the other hand, if the sequence $\{x_i\}$ is such that there are infinitely many terms x_n satisfying the inequality,

$$x_n < -M,$$

where M is any arbitrarily large positive number, then the lower limit l is $-\infty$.

For example, the sequence

(*a*) $$1, 2, 3, \cdots$$

has $L = +\infty$, whereas the sequence

(*b*) $$-1, -2, -3, \cdots$$

has $l = -\infty$. The sequence

(*c*) $$-1, 2, -3, 4, -5, 6, \cdots$$

has $L = +\infty$ and $l = -\infty$.

Some further examples of upper and lower limits of sequence may prove useful. The sequence

(*d*) $$-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, \cdots, \frac{(-1)^n}{n}, \cdots$$

has $L = 0$ and $l = 0$, but the sequence

$$+1, -1, +1, -1, +1, \cdots$$

has $l = -1$ and $L = 1$. The sequence

$$1, 2, 1/3, 4, \cdots, n^{(-1)^n}, \cdots$$

has $l = 0, L = \infty$, whereas

$$-1, -2, -1/3, -4, \cdots, -n^{(-1)^n}, \cdots$$

has $l = -\infty$ and $L = 0$.

It will be observed that no definition of the lower limit was supplied for sequences of the type (*a*), whereas sequences of the type (*b*) lack a definition of the upper limit. It is desirable to fill these gaps so that one can say that every sequence has uniquely defined upper and lower limits. Accordingly, it is agreed that if the sequence $\{x_i\}$ is such that $+\infty$ is its only limiting point, then the lower limit l of such a sequence is also $+\infty$. On the other hand, if $-\infty$ is the only limiting point of the sequence, then both the upper and lower limits are defined to be $-\infty$.

The sequence (*d*) is interesting because it has $L = l = 0$. There is a theorem in regard to bounded sequences whose upper and lower limits are equal.

Theorem. *A necessary and sufficient condition for the convergence of a bounded sequence is that the upper limit L be equal to the lower limit l.*

The proof of the theorem is not difficult. For, if

$$L = l = A,$$

then for any $\epsilon > 0$ there is at most a finite number of terms of $\{x_n\}$ such that

$$x_n > L + \epsilon \equiv A + \epsilon,$$

and also at most a finite number of x_n for which

$$x_n < l - \epsilon \equiv A - \epsilon.$$

Thus,

$$A - \epsilon < x_n < A + \epsilon, \qquad \text{if} \qquad n \geq p,$$

or

$$|x_n - A| < \epsilon, \qquad \text{if} \qquad n \geq p,$$

which is the statement that $\lim\limits_{n\to\infty} x_n = A$.

Conversely, if

$$\lim_{n\to\infty} x_n = A,$$

then for any $\epsilon > 0$ and for $n \geq p$

$$|x_n - A| < \epsilon \tag{7-3}$$

or

$$A - \epsilon < x_n < A + \epsilon.$$

Therefore,

$$x_n < A + \epsilon \tag{7-4}$$

for an infinite number of x_n, and the inequality

$$x_n < A - \epsilon$$

is true for at most a finite number of the x_n. Thus

$$l = A. \tag{7-5}$$

But (7-3) also states that

$$x_n > A - \epsilon$$

is true for an infinite number of x_n, and that there are at most a finite number of x_n for which

$$x_n > A + \epsilon.$$

Consequently,

$$L = A. \tag{7-6}$$

It follows from (7-5) and (7-6) that

$$l = L.$$

PROBLEMS

1. Which of the following sequences diverge?

(*a*) $0, 1, 0, 2, 0, 3, 0, 4, \cdots$;
(*b*) $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots$;
(*c*) $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \cdots$;
(*d*) $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \cdots$;
(*e*) $1, \frac{3}{2}, \frac{7}{2^2}, \cdots, \frac{2^{n+1} - 1}{2^n}, \cdots$;
(*f*) $\log 1, \log 2, \cdots, \log n, \cdots$;
(*g*) $\frac{1}{\log 2}, \frac{1}{\log 3}, \cdots, \frac{1}{\log n}, \cdots$;
(*h*) $\log \frac{1}{2}, \log \frac{1}{3}, \cdots, \log \frac{1}{n}, \cdots$.

2. Find the upper and lower limits for the sequences of Prob. 1.

8. Functions of a Single Variable. Let a pair of variables y and x be connected by some functional relation. The functional dependence of y upon x will be denoted by the symbol $y = f(x)$, which means that to every value of x under consideration there corresponds at least one value of y. There is no implication in the foregoing statement that there is an analytical expression or a formula connecting y with x. In the study of elementary analysis one usually thinks of a function as being given by a formula indicating the result of performing some operations on x, such as x^2, $\sqrt{1 - x}$, $\log x$, $\sin 2x$, etc. In the definition given here the term *function* has a much broader meaning.

It will be assumed that the variable x is a real variable and that the corresponding values of y are also real. The function $y = f(x)$ is called a *single-valued function* if to every value of x under consideration there corresponds one and only one value of y. Otherwise, the function $f(x)$ is called *multiple valued*. Thus $y = x^2$ is a single-valued function, whereas $y = \sin^{-1} x$ is a multiple-valued function of x.

It is frequently necessary to restrict the range of values which x is permitted to assume. The function $y = \sqrt{x}$ furnishes real values for y only if x is restricted to be positive or zero. It is convenient to think of the values of x as being associated with the points of the x-axis of a cartesian system of coordinates, so that one can speak of the interval consisting of all real numbers between a pair of given values (say, $x = a$ and $x = b$). Such an interval is denoted by the symbol (a, b), where $a < b$. If the end points of the interval (a, b) are excluded from consideration, the interval is said to be *open*, and the fact that x is not permitted to assume the values a and b is denoted by writing

$$a < x < b.$$

On the other hand, if the interval is closed, one writes

$$a \leq x \leq b.$$

The interval may be open at one end point and closed at the other. For example,

$$a \leq x < b$$

means that the interval is closed on the left, so that the point $x = a$ is included in the considerations, whereas $x = b$ is excluded.

If for all values of x in a given interval, $f(x)$ is never greater than some fixed number M, the number M is called an *upper bound* for $f(x)$ in that interval, whereas if $f(x)$ is never less than some number m, the latter is called a *lower bound*. One or both of these bounds may not exist. Thus $y = \sin x$ in $(0, 2\pi)$ has an upper bound $+1$ and a lower bound -1, whereas $y = \frac{1}{x}$ in the interval $(0, 1)$ has a lower bound unity and no upper bound. The function $y = \frac{1}{x} - x$ has neither upper nor lower bound in the interval $(0, \infty)$.

In considering the functional dependence of y upon x, it may happen that as x approaches, by any conceivable sequence of steps, the limit x_0, y also approaches a limit L, so that one can write

$$\lim_{x \to x_0} y = \lim_{x \to x_0} f(x) = L \tag{8-1}$$

The precise meaning of this symbolic expression (8-1) is embodied in the following definition of the limit of $f(x)$:

The function $f(x)$ approaches the limit L as x tends to limit *x_0 when, corresponding to any given positive number ϵ, one can find a number δ such that $|f(x) - L| < \epsilon$ for all values of x for which*

$$0 < |x - x_0| < \delta.$$

It should be noted carefully that the statement

$$0 < |x - x_0| < \delta$$

excludes the point $x = x_0$ from consideration, so that the numerical value of the difference must be less than ϵ for all points in the interval $(x_0 - \delta, x_0 + \delta)$, except at the point $x = x_0$. Moreover, no restriction is imposed as to the way in which x is permitted to approach x_0, so that the inequality

$$|f(x) - L| < \epsilon$$

must be satisfied whenever $x \to x_0$ through a set of values greater than or less than x_0.

It may be remarked that, since

$$\lim_{x \to x_0} f(x) = L \tag{8-1}$$

means $|f(x) - L| < \epsilon$ for all values of x for which

$$0 < |x - x_0| < \delta,$$

one can write (8-1) in an alternative form

$$f(x) = L + \eta, \tag{8-2}$$

where $|\eta| < \epsilon$ for all values of x such that $0 < |x - x_0| < \delta$. The particular form (8-2) of the definition (8-1) will be of frequent use in the following pages.

It is sometimes important to know the behavior of $f(x)$ when $x \to x_0$ in such a way that x always remains greater (or less) than x_0, that is, when x is allowed to approach x_0 from the right (or left) side only. In such cases it is convenient to adopt the following notation. If $x \to x_0$ from the left (that is, x is always less than x_0), one writes

$$x \to x_0-,$$

and the fact that $x \to x_0$ from the right (so that x is always greater than x_0) is denoted by

$$x \to x_0+.$$

These limits will be called the *left-hand* and the *right-hand limits*, respectively.

Also

$$\lim_{x \to x_0-} f(x) = a \qquad \text{or} \qquad f(x_0-) = a,$$

and

$$\lim_{x \to x_0+} f(x) = b \qquad \text{or} \qquad f(x_0+) = b,$$

will stand for the left-hand and right-hand limits of $f(x)$. The symbols $f(x_0+)$ and $f(x_0-)$ should not be confused with the symbol $f(x_0)$, which means the value of the function $f(x)$ calculated at the point $x = x_0$. In fact $f(x_0+)$ and $f(x_0-)$ may exist, whereas the function $f(x)$ may not be defined at $x = x_0$. To clarify this, consider an example. Let $f(x) = \dfrac{1}{1 - e^{-\frac{1}{x}}}$. It is readily established that

$$\lim_{x \to 0-} f(x) \equiv f(0-) = 0,$$

and

$$\lim_{x \to x_0+} f(x) \equiv f(0+) = 1,$$

while

$$f(0) = \frac{1}{1 - e^{\frac{1}{0}}},$$

which is utterly devoid of sense since division by zero is undefined (Fig. 5).

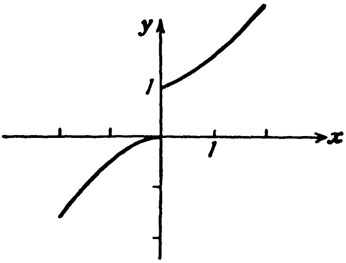

FIG. 5.

Since another example may prove useful, let $f(x) = 2^{\frac{1}{x-1}}$. In this case (see Fig. 6) $f(1+) = \infty$, $f(1-) = 0$, but $f(1) = 2^{1/0}$ is meaningless.

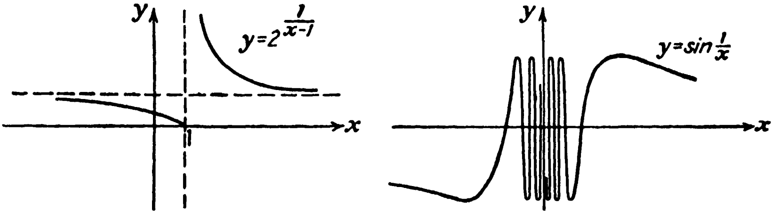

FIG. 6.

The function $y = \sin\frac{1}{x}$ has neither $f(0+)$ nor $f(0-)$, since $\sin\frac{1}{x}$ oscillates between its upper and lower bounds $+1$ and -1 as $x \to 0$ from either side (Fig. 6).

PROBLEMS

✓1. Find the right-hand and left-hand limits of $f(x) = x^2$ as $x \to 2$. What is $\lim_{x \to 2} x^2$?

✓2. Find $f(0+)$ and $f(0-)$ if $f(x) = \dfrac{e^{1/x}}{1 + e^{1/x}}$. What is $\lim_{x \to 0} f(x)$?

3. What are $f(0+)$ and $f(0-)$ if $f(x) = x \sin\frac{1}{x}$? Sketch the graph of $f(x)$ in the vicinity of $x = 0$. Find $\lim_{x \to 0} f(x)$.

4. Find $f(2+)$ and $f(2-)$, if $f(x) = \dfrac{1}{x-2}$.

5. What is $\lim\limits_{x\to\frac{\pi}{6}} \sin x$? $\sin\left(\dfrac{\pi}{6}+\right)$? $\sin\left(\dfrac{\pi}{6}-\right)$?

6. Find $\lim\limits_{x\to 0} \dfrac{x}{1-e^{1/x}}$.

Sketch the graph of the function in the neighborhood of $x = 0$.

9. Theorems on Limits.

Theorem 1. *The limit of a sum is equal to the sum of the limits.* If $\lim\limits_{x\to x_0} f_1(x) = L_1$ and $\lim\limits_{x\to x_0} f_2(x) = L_2$, then

$$\lim_{x\to x_0} [f_1(x) + f_2(x)] = L_1 + L_2.$$

The proof of this theorem follows directly from the definition of the limit. It is required to show that for any preassigned positive number ϵ one can determine a number δ such that

$$|f_1(x) + f_2(x) - L_1 - L_2| < \epsilon,$$

whenever x lies in the interval $0 < |x - x_0| < \delta$. But, by hypothesis, $\lim\limits_{x\to x_0} f_1(x) = L_1$, so that

$$(9\text{-}1) \qquad |f_1(x) - L_1| < \frac{\epsilon}{2}, \qquad \text{whenever} \qquad 0 < |x - x_0| < \delta_1.$$

Similarly,

$$(9\text{-}2) \qquad |f_2(x) - L_2| < \frac{\epsilon}{2}, \qquad \text{whenever} \qquad 0 < |x - x_0| < \delta_2.$$

Now if δ is chosen to be the smaller of δ_1 and δ_2, then it follows from (9-1) and (9-2) that*

$$|f_1(x) + f_2(x) - L_1 - L_2| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

whenever $0 < |x - x_0| < \delta$.

---

* Since the sum of the absolute values is not less than the absolute value of the sum.

But this is precisely the statement that

$$\lim_{x\to x_0} [f_1(x) + f_2(x)] = L_1 + L_2.$$

The extension to a greater number of functions follows immediately.

Theorem 2. *The limit of a product is equal to the product of the limits.*

If $\lim_{x\to x_0} f_1(x) = L_1$ and $\lim_{x\to x_0} f_2(x) = L_2$, then

$$\lim_{x\to x_0} f_1(x)f_2(x) = L_1L_2. \tag{9-3}$$

It is required to show that

$$|f_1(x)f_2(x) - L_1L_2| < \epsilon, \qquad \text{whenever} \qquad 0 < |x - x_0| < \delta.$$

But

$$\begin{aligned} |f_1(x)f_2(x) - L_1L_2| &= |f_1(x)f_2(x) - L_1f_2(x) + L_1f_2(x) - L_1L_2| \\ &\leq |f_1(x)f_2(x) - L_1f_2(x)| + |L_1f_2(x) - L_1L_2| \\ &= |f_2(x)| \cdot |f_1(x) - L_1| + |L_1| \cdot |f_2(x) - L_2|. \end{aligned}$$

By hypothesis $\lim_{x\to x_0} f_2(x) = L_2$, so that $f_2(x)$ is surely bounded in the neighborhood of the point $x = x_0$. Hence, $|f_2(x)| < M$ for all values of x such that $0 < |x - x_0| < \delta'$. Then

$$|f_1(x)f_2(x) - L_1L_2| < M \cdot |f_1(x) - L_1| + |L_1| \cdot |f_2(x) - L_2|.$$

Moreover, since $\lim_{x\to x_0} f_1(x) = L_1$ and $\lim_{x\to x_0} f_2(x) = L_2$, corresponding to any $\epsilon > 0$, one can find a positive number $\delta < \delta'$ such that

$$|f_1(x) - L_1| < \frac{\epsilon}{2M},$$

and

$$|f_2(x) - L_2| < \frac{\epsilon}{2|L_1|},$$

whenever $0 < |x - x_0| < \delta$. Thus

$$|f_1(x)f_2(x) - L_1L_2| < M \cdot \frac{\epsilon}{2M} + |L_1| \cdot \frac{\epsilon}{2|L_1|} = \epsilon,$$

whenever $0 < |x - x_0| < \delta$, which is equivalent to (9-3).

The result can be immediately generalized to any number of functions.

Theorem 3. *The limit of a quotient is equal to the quotient of the limits, provided the limit of the denominator is not zero.*

If $\lim_{x\to x_0} f_1(x) = L_1$ and $\lim_{x\to x_0} f_2(x) = L_2 \neq 0$, then it must be proved that

$$\lim_{x\to x_0} \frac{f_1(x)}{f_2(x)} = \frac{L_1}{L_2}.$$

Note that

$$\begin{aligned}\frac{f_1(x)}{f_2(x)} - \frac{L_1}{L_2} &= \left[\frac{f_1(x)}{L_2} - \frac{L_1}{L_2}\right] + \left[-\frac{f_1(x)}{L_2} + \frac{f_1(x)}{f_2(x)}\right] \\ &= \left[\frac{1}{L_2}[f_1(x) - L_1]\right] + \left[\frac{f_1(x)}{L_2 f_2(x)}[L_2 - f_2(x)]\right].\end{aligned}$$

But, since the limits of $f_1(x)$ and $f_2(x)$ exist and $\lim_{x\to x_0} f_2(x) \neq 0$, one can certainly make the numerical value of each bracketed term of the right-hand member of the last expression less than $\frac{\epsilon}{2}$, when $0 < |x - x_0| < \delta$. Therefore,

$$\left|\frac{f_1(x)}{f_2(x)} - \frac{L_1}{L_2}\right| < \epsilon, \qquad \text{when} \qquad 0 < |x - x_0| < \delta,$$

which is the required result.

PROBLEMS

1. Give a careful justification of the inequality

$$|f_2(x) - L_2| \cdot \left|\frac{f_1(x)}{L_2 f_2(x)}\right| < \frac{\epsilon}{2}, \qquad \text{when} \qquad 0 < |x - x_0| < \delta,$$

used in the proof of Theorem 3.

2. If n is a positive integer, prove that $\lim_{x\to 0} x^n = 0$, and, hence, that

$$\lim_{x\to 0} (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n) = a_0.$$

10. The Base of the Natural Logarithms. An interesting application of the criterion for convergence of monotone sequences to the determination of the limiting value of the function $\left(1 + \frac{1}{x}\right)^x$ as $x \to \infty$, is contained in this section.

It will be shown first that $\left(1 + \frac{1}{x}\right)^x$ approaches a limit when x tends to infinity by assuming positive integral values. In this case the problem reduces to the study of the behavior of the sequence $\{x_n\}$, the nth term of which is given by the formula

$$x_n = \left(1 + \frac{1}{n}\right)^n.$$

It will be shown that the sequence $\{x_n\}$ is monotone increasing, bounded, and hence convergent.

Now the ratio of x_n to x_{n-1} is

$$\begin{aligned}\frac{x_n}{x_{n-1}} = \frac{\left(1 + \frac{1}{n}\right)^n}{\left(1 + \frac{1}{n-1}\right)^{n-1}} &= \left(\frac{1 + \frac{1}{n}}{1 + \frac{1}{n-1}}\right)^n \cdot \left(1 + \frac{1}{n-1}\right)\\ &= \left(\frac{n^2 - 1}{n^2}\right)^n \cdot \frac{n}{n-1}\\ &= \left(1 - \frac{1}{n^2}\right)^n \cdot \frac{n}{n-1}.\end{aligned}$$

But by the inequality of Bernoulli,* for any $n > 1$,

$$\left(1 - \frac{1}{n^2}\right)^n > 1 - \frac{n}{n^2} = \frac{n-1}{n}.$$

Hence,

$$\frac{x_n}{x_{n-1}} > 1,$$

so that the sequence $\{x_n\}$ is monotone increasing. Moreover it is bounded on the right, since

$$\left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n}\right)^{n+1}$$

and the sequence $\left(1 + \frac{1}{n}\right)^{n+1}$ is monotone decreasing.†

---

* See Sec. 4.

† See Prob. 2, p. 31.

Therefore,

$$\left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n}\right)^{n+1} \le (1+1)^2 = 4,$$

and thus $\left(1+\frac{1}{n}\right)^n$ converges to some limit, which will be denoted by the letter e. Hence,

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e. \tag{10-1}$$

Consider now the function

$$\left(1+\frac{1}{x}\right)^x,$$

where x is allowed to take on any sequence of positive real values tending to infinity. Let n be the greatest integer contained in any given value of x, so that

$$n \le x < n+1.$$

If x tends to infinity, the number n does likewise, and one can write

$$\left(1+\frac{1}{n+1}\right)^n < \left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{n}\right)^{n+1} \tag{10-2}$$

But it follows from (10-1) and from Sec. 9 that

$$\lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^n = \lim_{n\to\infty}\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{1+\frac{1}{n+1}} = e.$$

Also

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1} = \lim_{n\to\infty}\left[\left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right)\right] = e.$$

Since the extreme members of the inequality (10-2) tend to the same limit, it follows that

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e. \tag{10-3}$$

The problem of computing a rational approximation to the number e with any desired degree of accuracy is discussed in Chap. IX.

PROBLEMS

1. Show, with the aid of the result (10-3), that

$$\lim_{x \to -\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Hint: Introduce the new variable $\xi \equiv -(1 + x)$ and let ξ tend to infinity through any set of positive values.

2. Show that the sequence $\{y_n\}$, the nth term of which is

$$y_n = \left(1 + \frac{1}{n}\right)^{n+1}$$

is monotone decreasing and bounded. Form the sequence $\{y_n - x_n\}$ where $\{x_n\}$ is the sequence discussed in Sec. 10, and hence, deduce that $\lim_{n \to \infty} y_n = e$. Show that $1 \leq e < 3$.

11. Continuity. In defining the limit of $f(x)$ as $x \to x_0$, no assumption was made regarding the value of the function at $x = x_0$. In fact, it was emphasized that the function need not be defined at $x = x_0$.

The function $f(x)$ *is said to be continuous at the point* $x = x_0$, *if the right-hand and left-hand limits of* $f(x)$ *as* $x \to x_0$ *are equal and finite and coincide with the value of the function at* $x = x_0$.

This definition can be stated compactly in the form of an equation

$$\lim_{x \to x_0} f(x) = f(x_0). \tag{11-1}$$

A more explicit way of restating the definition of continuity (11-1) is the following:

If for any preassigned positive number ϵ, *one can find a positive number* δ *such that*

$$|f(x) - f(x_0)| < \epsilon, \qquad \textit{whenever} \qquad |x - x_0| < \delta, \tag{11-2}$$

then the function $f(x)$ *is continuous at the point* $x = x_0$.

A geometrical interpretation of the statement embodied in (11-2) is immediate. Corresponding to any preassigned positive number ϵ, there can be determined an interval of width 2δ about the point $x = x_0$ (Fig. 7) such that, for every x lying in the interval $(x_0 - \delta,\ x_0 + \delta)$, $f(x)$ is confined to lie between $f(x_0) + \epsilon$ and $f(x_0) - \epsilon$.

An alternative way of writing the defining equation (11-1) is frequently useful. Thus (11-1) can be written in the form

$$f(x) = f(x_0) + \eta, \tag{11-3}$$

where $\lim\limits_{x \to x_0} \eta = 0$. The identity of (11-1) and (11-3) follows immediately upon taking the limit of both members of (11-3).

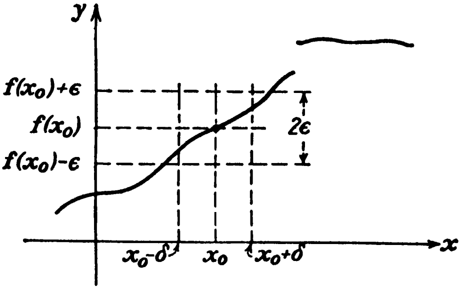

FIG. 7.

Two important remarks are in order:

a. If the function is continuous, then $\lim\limits_{x \to x_0} f(x)$ must exist,

b. The function must be defined at the point of continuity.

The significance of these remarks is made clear by the following examples:

Illustrative Example 1. To test the continuity of $f(x) = x^2$ at $x = 2$, note that $f(2) = 4$. Moreover, $\lim\limits_{x \to 2} x^2 = 4$. Hence, $f(x) = x^2$ is continuous at $x = 2$.

Illustrative Example 2. Let $f(x) = \dfrac{x^2 - 9}{x - 3}$. Now

$$f(3) = \frac{3^2 - 9}{3 - 3} = \frac{0}{0},$$

which is meaningless. But $\lim\limits_{x \to 3} \dfrac{x^2 - 9}{x - 3} = \lim\limits_{x \to 3} (x + 3) = 6$. Hence, if $f(x)$ is defined to be equal to 6 at $x = 3$, then $f(x)$ will be continuous at $x = 3$, but not otherwise.

Illustrative Example 3. Let $f(x) = \dfrac{\sin x}{x}$ if $x \neq 0$, and let $f(x) = 1$ if $x = 0$. Now $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1$, which coincides with the definition of $f(x)$ at $x = 0$, so that $f(x)$ is continuous at $x = 0$.

If it happens that the function $f(x)$ is continuous at every point of the interval (a, b), then the function is said to be *continuous over the interval* (a, b). If the interval is closed, the continuity at the end points $x = a$ and $x = b$ is defined by the equations

$$\lim_{x \to a+} f(x) = f(a) \qquad \text{and} \qquad \lim_{x \to b-} f(x) = f(b),$$

since such a symbol as

$$\lim_{x \to a} f(x)$$

has no meaning if the consideration of functional values is restricted to the closed interval (a, b).

It is clear from the foregoing that any function that can be represented graphically as an unbroken curve is continuous. However, not every continuous function can be represented graphically. For example, let

$$f(x) = x \sin \frac{1}{x},$$

and consider the behavior of the function in the vicinity of $x = 0$. Inasmuch as $\sin \dfrac{1}{x}$ is defined at every point of the x-axis, except $x = 0$, one can calculate the values of $f(x)$ for any sequence of values of x converging to zero. Moreover, since $\sin \dfrac{1}{x}$ is bounded,

$$\lim_{x \to 0} x \sin \frac{1}{x} = 0.$$

At the point $x = 0$, $f(x)$ is undefined, but if it is agreed to assign the value zero to $f(0)$, the function

$$\begin{aligned} f(x) &= x \sin \frac{1}{x}, \qquad && \text{if} \qquad x \neq 0, \\ &= 0, && \text{if} \qquad x = 0, \end{aligned}$$

will be continuous at $x = 0$; yet the graph of the function in the vicinity of the origin cannot be drawn, since the function oscillates infinitely often in any interval containing the origin (Fig. 8).

Whenever a function fails to be continuous at a point, it is called *discontinuous* at that point. The discontinuity at $x = x_0$ is called *infinite* whenever $f(x)$ becomes infinite as $x \to x_0$. If the function is discontinuous at a point $x = x_0$ but is bounded in the vicinity of that point, the discontinuity is called *finite* or *ordinary*. If $f(x_0+)$ and $f(x_0-)$ exist and are unequal, their difference

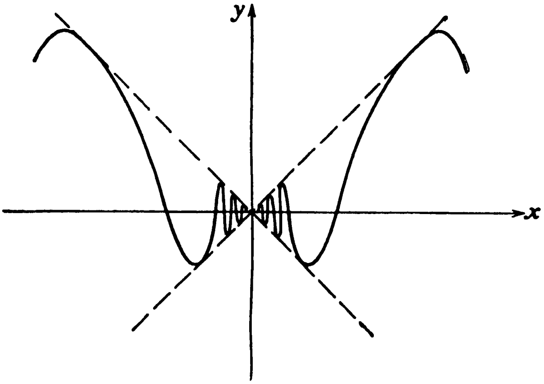

Fig. 8.—Graph of $y = x \sin \frac{1}{x}$.

$$f(x_0+) - f(x_0-)$$

is called the *jump in the function.* Thus the function

$$f(x) = \frac{1}{1 - e^{-\frac{1}{x}}}$$

discussed in Sec. 8 has a jump of one unit at $x = 0$, whereas $f(x) = 2^{\frac{1}{x-1}}$ has an infinite discontinuity at $x = 1$.

A function having a finite number of jumps in a given interval is called *piece-wise continuous* or *sectionally continuous.*

PROBLEM

Discuss the continuity of the following functions:

(a) $y = x^2 - 2x$;

(b) $y = 2^x$;

(c) $y = e^{\frac{1}{x}}$;

(d) $y = \log x$;

(e) $y = x \sin \frac{\pi}{x}$;

(f) $y = \sin \frac{\pi}{x}$;

(g) $y = \frac{1}{(1-x)^2}$;

(h) $y = \frac{1}{1-x^2}$;

(i) $y = \frac{1}{x^2 - 3x + 2}$;

(j) $y = x + \frac{1}{x}$;

(k) $y = \cot x$;

(l) $y = \frac{1 - \cos x}{x}$, if $x \neq 0$,
$= 0$, if $x = 0$;

(m) $y = \frac{x^2 - 1}{x - 1}$, if $x \neq 1$,
$= 2$, if $x = 1$;

(n) $y = e^{-\frac{1}{x^2}}$, if $x \neq 0$,
$= 0$, if $x = 0$;

(o) $y = 1 - x$, if $x > 0$,
$= 1 + x$, if $x < 0$,
$= 1$, if $x = 0$.

12. Properties of Continuous Functions. It was seen from the discussion of $f(x) = x \sin \frac{1}{x}$ in Sec. 11 that the intuitive idea of a continuous function as a function that can be represented graphically as an unbroken curve is quite inadequate to account for functions that are continuous according to the definition given above and that may not be capable of graphical representation in the neighborhood of a point of continuity. A function that can be represented graphically as an unbroken curve is, of course, continuous.

The seemingly simple concept of continuity becomes, on careful examination, exceedingly complex. A detailed study of the properties of continuous functions is the province of the theory of functions of a real variable, but some of the simpler theorems follow directly from the definition of a continuous function and upon the application of the theorems on limits.

Theorem 1. *The sum, difference, product, and quotient of any finite number of functions, each of which is continuous at a given point $x = x_0$, will be continuous at $x = x_0$, provided that the denominator of the quotient does not vanish at $x = x_0$.*

The proof of this theorem follows immediately from the definition of continuity with the aid of the theorems of Sec. 9. It is left as an exercise for the reader.

Theorem 2. *If* $y = f_1(x)$ *is a continuous function at the point* $x = x_0$, *and if* $z = f_2(y)$ *is a continuous function at* $y = y_0$, *where* $y_0 = f_1(x_0)$, *then the composite function* $z = f_2[f_1(x)]$ *is continuous at* $x = x_0$.

By hypothesis $z = f_2(y)$ is continuous at $y = y_0$, so that

$$\lim_{y \to y_0} f_2(y) = f_2(y_0).$$

Hence, corresponding to any $\epsilon > 0$, one can find a positive number δ_1 such that

$$|f_2(y) - f_2(y_0)| < \epsilon, \qquad \text{when} \qquad |y - y_0| < \delta_1.$$

But

$$f_2(y) = f_2[f_1(x)],$$

and, therefore, the foregoing inequality can be written as

$$(12\text{-}1) \quad |f_2[f_1(x)] - f_2(y_0)| < \epsilon, \qquad \text{when} \qquad |f_1(x) - y_0| < \delta_1.$$

But $f_1(x)$ is also continuous at $x = x_0$; hence, corresponding to the positive number δ_1, one can find a positive number δ such that

$$|f_1(x) - y_0| < \delta_1, \qquad \text{whenever} \qquad |x - x_0| < \delta.$$

Thus (12-1) can be rewritten to read

$$|f_2[f_1(x)] - f_2(y_0)| < \epsilon, \qquad \text{whenever} \qquad |x - x_0| < \delta,$$

which is another way of saying that

$$\lim_{x \to x_0} f_2[f_1(x)] = f_2(y_0).$$

The latter statement is precisely the definition of continuity.

Theorem 3. *If* $f(x)$ *is continuous at* $x = x_0$ *and* $f(x_0) \neq 0$, *then one can find an interval about the point* $x = x_0$ *such that* $f(x)$ *has the same sign as* $f(x_0)$ *throughout this interval.*

The continuity of $f(x)$ at $x = x_0$ means that corresponding to any $\epsilon > 0$ one can find a number δ such that

$$(12\text{-}2) \quad |f(x) - f(x_0)| < \epsilon, \qquad \text{when} \qquad |x - x_0| < \delta.$$

Now (12-2) can be written as

$$(12\text{-}3) \quad f(x_0) - \epsilon < f(x) < f(x_0) + \epsilon, \qquad \text{when} \qquad |x - x_0| < \delta,$$

and if ϵ is chosen to be equal to $|f(x_0)|$, the inequality (12-3) reads

$$(12\text{-}4) \qquad f(x_0) - |f(x_0)| < f(x) < f(x_0) + |f(x_0)|,$$

when $|x - x_0| < \delta_0$.
Hence, if $f(x_0) > 0$, (12-4) becomes

$$0 < f(x) < 2f(x_0),$$

so that $f(x)$ is positive for all values of x in the interval

$$x_0 - \delta_0 < x < x_0 + \delta_0.$$

On the other hand, if $f(x_0) < 0$, then the inequality (12-4) shows that $f(x) < 0$ in the interval $x_0 - \delta_0 < x < x_0 + \delta_0$.

The reader might be inclined to regard the theorems just established as obvious from graphical considerations, which they indeed are if one restricts oneself to those continuous functions which can be represented by unbroken curves.

13. Uniform Continuity. In defining the continuity of a function at a point $x = x_0$, it was stated that, corresponding to an arbitrary positive number ϵ, one can find a positive number δ (depending on ϵ) such that for all values of x in the interval of width 2δ, with x_0 as the midpoint of the interval, the numerical value of the difference $f(x) - f(x_0)$ will be less than ϵ.

It is important to observe that the magnitude of δ depends on both the number ϵ and the value of x under consideration. This can be made clear by considering the function $f(x)$ whose graph is shown in Fig. 9. Let ϵ be prescribed; then for all values of x in the interval of width $2\delta_1$ about $x = x_1$

$$|f(x) - f(x_1)| < \epsilon,$$

whereas in order to make

$$|f(x) - f(x_2)| < \epsilon$$

one would require the values of x in the neighborhood of the point $x = x_2$ to lie in a much smaller interval of width $2\delta_2$. Now the remarkable fact about functions continuous in *a closed interval* (a, b) is that one can find a number δ which is independent of the particular value of x under consideration and thus depends on ϵ alone.* Or, to put it differently, if the interval (a, b) is divided in any manner whatever into a finite number of subintervals of widths not exceeding 2δ, then for any two points in the same subinterval, the numerical value of the difference between the values of $f(x)$ will be less than 2ϵ. To signify the fact that a single number δ will serve *for all values of x in the interval* (a, b), one says that the function is *uniformly continuous* in the interval.

---

* The proof of this is omitted. See E. W. Hobson, Theory of Functions of a Real Variable, p. 290; E. Goursat, Cours d'analyse mathématique, 5th ed., vol. 1, p. 14; C. J. de la Vallee Poussin, Cours d'analyse infinitesimale, vol. 1, p. 58.

It is not difficult to see that a function which is continuous in the open interval (a, b) need not be uniformly continuous.

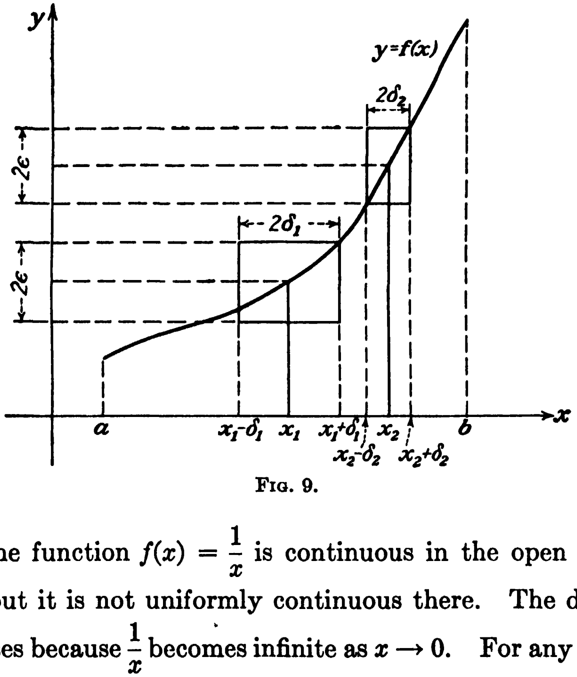

FIG. 9.

Thus, the function $f(x) = \frac{1}{x}$ is continuous in the open interval $(0, 1)$, but it is not uniformly continuous there. The difficulty here arises because $\frac{1}{x}$ becomes infinite as $x \to 0$. For any nonzero value of x the function is continuous, but it is impossible to subdivide the interval $(0, 1)$ into a finite number of subintervals such that the numerical value of the difference of the functional values, for any pair of points in each of the partial intervals, is less than a preassigned number ϵ.

The property of uniform continuity of functions defined over a closed interval is used in the following theorems.*

---

* The proofs of Theorems 1 and 2 will be found in E. W. Hobson, Theory of Functions of a Real Variable, pp. 281–285; E. Goursat, Cours d'analyse

Theorem 1. *A function $f(x)$ continuous in a closed interval (a, b) is bounded in that interval.*

The theorem states that one can find a positive number M such that $|f(x)| \leq M$, for all values of x in the interval $a \leq x \leq b$.

Theorem 2. *If $f(x)$ is continuous in a closed interval (a, b), then there exists in this interval at least one value of x for which $f(x)$ takes its maximum value and at least one value of x for which it assumes its minimum value.*

Theorem 3. *If $f(x)$ is continuous in a closed interval (a, b) then, as x takes all values between a and b, the function $f(x)$ takes at least once all values intermediate to $f(a)$ and $f(b)$.*

Corollary. *If $f(a)$ and $f(b)$ are of opposite signs then there exists at least one point intermediate to a and b such that $f(x)$ vanishes at that point.*

In order to give an idea of the type of reasoning employed in proving these theorems, the proof of Theorem 3 follows. The corollary to Theorem 3 will be established first.

Let the signs of $f(a)$ and $f(b)$ be opposite and set $\frac{1}{2}(a + b) = x_1$. If $f(x_1)$ vanishes, the corollary is demonstrated. If $f(x_1) \neq 0$, let (a_1, b_1) be that one of the two equal intervals (a, x_1) and (x_1, b) for which $f(a_1)$ and $f(b_1)$ are of opposite signs. Form

$$x_2 = \tfrac{1}{2}(a_1 + b_1).$$

If $f(x_2) = 0$, the corollary is established. If $f(x_2) \neq 0$, let (a_2, b_2) be that one of the two equal intervals (a_1, x_2) and (x_2, b_1) for which $f(a_2)$ and $f(b_2)$ are of opposite signs. If in continuing this process one does not obtain a number x_i such that $f(x_i) = 0$, there results a sequence of intervals

$$(a, b), (a_1, b_1), (a_2, b_2), \cdots, (a_n, b_n), \cdots,$$

such that each of them is contained in the preceding, and such that the function $f(x)$ has opposite signs at the end points of these intervals.

The numbers $a, a_1, a_2, \cdots, a_n, \cdots$ form a monotone-increasing sequence and remain less than b, so that they tend to some limit* c contained between a and b. The numbers $b, b_1, b_2, \cdots, b_n, \cdots$, on the other hand, are always greater than a

---

mathématique, 5th ed., vol. 1, p. 16 (the English translation of this book contains these proofs on p. 143).

* See Sec. C

and form a decreasing sequence, so that they also converge to some limit c'. The limits c and c' are equal since the difference

$$b_n - a_n = \frac{1}{2^n}(b - a)$$

tends to zero when n increases indefinitely.

But $f(x)$ is continuous in the interval (a, b), and therefore the difference

$$f(b_n) - f(a_n)$$

approaches zero when $n \to \infty$. Moreover, since the signs of $f(b_n)$ and $f(a_n)$ are opposite,

$$\lim_{n \to \infty} |f(b_n) - f(a_n)| = 2|f(c)| = 0,$$

or

$$f(c) = 0.$$

Thus the corollary to Theorem 3 is demonstrated.

In order to establish the theorem itself, let k be some number between $f(a)$ and $f(b)$ and form the function

$$F(x) = f(x) - k.$$

The function $F(x)$ is continuous in the interval (a, b) and has opposite signs at the end points of the interval so that there exists at least one value c intermediate to a and b such that $F(c) = 0$. But, if $F(c) = 0$,

$$f(c) = k,$$

and the theorem is demonstrated.

It may be remarked that if $f(x)$ is steadily increasing or decreasing and is continuous in the closed interval (a, b), then there will be exactly one value of x, intermediate to a and b, for which $f(x)$ assumes a prescribed value k, intermediate to $f(a)$ and $f(b)$.

CHAPTER II

DERIVATIVES AND DIFFERENTIALS

14. Derivatives.*

Definition. *If* $x = x_0$ *is any point of the interval* (a, b) *and if* $f(x)$ *is a function of* x *defined over* (a, b), *then*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

is called the derivative of $f(x)$ *at the point* $x = x_0$.

Denoting the derivative of $f(x)$ at $x = x_0$ by the symbol $f'(x_0)$ and recalling the definition of the limit† give

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) + \eta, \tag{14-1}$$

where

$$\lim_{h \to 0} \eta = 0.$$

Rewriting (14-1) gives

$$f(x_0 + h) - f(x_0) = hf'(x_0) + h\eta;$$

hence,

$$|f(x_0 + h) - f(x_0)| \leq |hf'(x_0)| + |h\eta|.$$

Now if any $\epsilon > 0$ is prescribed,

$$|hf'(x_0)| + |h\eta|$$

can be made less than ϵ by making h sufficiently small. Then, for any value of x in the interval $x_0 - h < x < x_0 + h$,

$$|f(x) - f(x_0)| < \epsilon,$$

---

* The reader is assumed to be familiar with calculations involving the use of the formulas for the differentiation of the so-called *elementary functions*. It is advisable to review them before proceeding with the study of this chapter.

† See p. 24.

which states that $f(x)$ is necessarily continuous at $x = x_0$ if it possesses the derivative at $x = x_0$. This result can be stated as a theorem:

Theorem. *Any function having the derivative at a point is necessarily continuous at that point.*

The converse of the theorem is not true, as can be seen from the following example (Fig. 10): let

$$\begin{aligned} f(x) &= -x + 1, &\quad &\text{when} &\quad &x > 0, \\ f(x) &= x + 1, &\quad &\text{when} &\quad &x \leq 0. \end{aligned}$$

The function is obviously continuous at $x = 0$, but the derivative at $x = 0$ does not exist since the limit of the difference quotient

$$\frac{f(x + h) - f(x)}{h},$$

independent of the mode of approach of h to zero, does not exist. This becomes clear immediately if one recalls the geometrical interpretation of the derivative as the slope of the tangent line to the locus of $f(x)$. If the point $x = 0$ is approached from the right, the slope of the tangent line is -1, whereas it is equal to $+1$ when $x \to 0$ from the left.

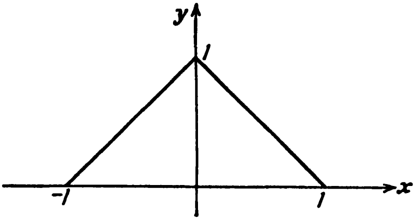

FIG. 10.—Graph of the function
$y = -x + 1, \quad \text{if} \quad x > 0;$
$ = x + 1, \quad \text{if} \quad x \leq 0.$

As a more interesting example consider the function

$$\begin{aligned} f(x) &= x \sin \frac{1}{x}, &\quad &\text{if} &\quad &x \neq 0, \\ &= 0, &\quad &\text{if} &\quad &x = 0, \end{aligned}$$

which was shown* to be continuous at $x = 0$. However, the difference quotient is

---

* See p. 33.

$$\frac{f(0+h)-f(0)}{h}=\frac{h\sin\frac{1}{h}-0}{h}=\sin\frac{1}{h},$$

and $\lim\limits_{h\to 0}\sin\frac{1}{h}$ does not exist, since $\sin\frac{1}{h}$ oscillates between $+1$ and -1 as $h \to 0$.

To be sure, this function has a derivative at every point except $x = 0$, but it is possible to construct functions that are continuous for all values of x in a given interval, but that do not have a derivative at a single point of the interval.* Functions that have derivatives are called *differentiable functions.*

PROBLEMS

1. Show from the definition that $f'(1) = \frac{1}{2}$, if $f(x) = \sqrt{x}$.

2. Does the derivative of

$$\begin{aligned} f(x) &= \frac{x-1}{1+e^{1/(x-1)}}, &&\text{if} \quad x \neq 1, \\ &= 0, &&\text{if} \quad x = 1, \end{aligned}$$

exist at $x = 1$?

3. Let

$$\begin{aligned} f(x) &= x^2 \sin\frac{1}{x}, &&\text{if} \quad x \neq 0, \\ &= 0, &&\text{if} \quad x = 0. \end{aligned}$$

Then

$$f'(0) = \lim_{h\to 0}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0}\frac{h^2\sin 1/h}{h} = 0.$$

For any point other than $x = 0$ a formula for the differentiation of the product can be used Thus,

$$f'(x) = 2\,x\sin\frac{1}{x} - \cos\frac{1}{x},$$

which is meaningless at $x = 0$. Why?

15. Differentials. If the functional relation between y and x is given in the form

$$y = f(x),$$

---

* K. Weierstrass was the first to demonstrate the existence of such functions in 1861, Many examples of such functions have been constructed since then.

the variables y and x are called, respectively, *dependent* and *independent variables.*

Let Δx be an arbitrary increment in x. The resulting change in the dependent variable y when x changes by an amount Δx will be written as Δy. The increment of the independent variable x will be named the differential of x and denoted by either Δx or dx. The differential of the dependent variable y, on the other hand, is defined by the formula

$$dy = f'(x)\,dx, \tag{15-1}$$

so that the differential dy and the increment Δy of the dependent variable, in general, are quite distinct. From the defining equation (15-1) of the differential, it is clear that the derivative of $y = f(x)$ can be expressed as the quotient of two differentials, namely,

$$f'(x) = \frac{dy}{dx}.$$

In this expression dx is assigned at pleasure, but dy is determined from (15-1).

From the definition of the derivative

$$f'(x) \equiv \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$

it is seen that

$$\frac{\Delta y}{\Delta x} = f'(x) + \eta,$$

where

$$\lim_{\Delta x \to 0} \eta = 0.$$

Solving for Δy gives

$$\Delta y = f'(x)\,\Delta x + \eta\,\Delta x$$

and, since $\Delta x \equiv dx$,

$$\Delta y - dy = \eta\,\Delta x.$$

For small values of Δx this difference is small. Geometrically $\eta\,\Delta x$ represents the portion of the increment Δy cut off by the tangent line to the locus of $f(x)$ at the point where the derivative is computed (see Fig. 11).

Example. Let $y = x^2$. Then $dy = 2x\,dx$, while

$$\Delta y = (x + \Delta x)^2 - x^2 = 2x\,\Delta x + (\Delta x)^2.$$

In this case $\eta\,\Delta x = (\Delta x)^2$, so that $\eta = \Delta x$.

PROBLEMS

1. Find dy and Δy, if $y = x^3$.
2. Find dy and Δy, if $y = \sqrt{x + 1}$, $x \geq 0$.
3. Find dy and Δy, if $y = \sin x$.

16. Derivatives of Composite Functions. If $z = f(y)$ and $y = \varphi(x)$, one can eliminate y between these relations and obtain

$$z = f[\varphi(x)] \equiv F(x).$$

The function $z = F(x)$ may have a derivative $\frac{dz}{dx}$, but often it is undesirable to perform the elimination of y in order to calculate $\frac{dz}{dx}$. If the functions $f(y)$ and $\varphi(x)$ are differentiable, it is possible to calculate $\frac{dz}{dx}$ without performing the elimination.

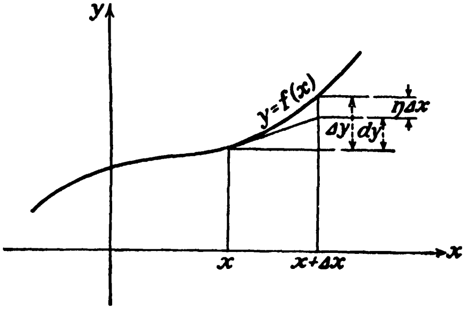

FIG. 11.

Note that if x is given an increment Δx, then y will acquire an increment Δy, and $\Delta y \to 0$ when $\Delta x \to 0$, since $\varphi(x)$ is assumed to be a differentiable and, therefore, continuous function. But corresponding to an increment Δy there will be an increment Δz. Thus one can write

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \cdot \frac{\Delta y}{\Delta x} \tag{16-1}$$

and

$$\lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta y} \cdot \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$

or

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

In writing (16-1) it is tacitly assumed that $\Delta y \neq 0$. There remains to be considered the case when $\Delta y = 0$ for some choice of the values of Δx approaching zero. Now

$$\Delta z = f(y + \Delta y) - f(y),$$

and if $\Delta y = 0$, then $\Delta z = 0$, so that

$$\frac{dz}{dx} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = 0.$$

Moreover, if $\Delta y = 0$ then

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 0.$$

Hence, the formula

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} \tag{16-2}$$

is true in this case also.

As an illustration of this case, consider the function

$$z = u^2$$

and

$$\begin{aligned} u &= x^2 \sin \frac{1}{x}, &\quad &\text{if} \quad x \neq 0, \\ &= 0, &\quad &\text{if} \quad x = 0. \end{aligned}$$

Then

$$\Delta z = 2u\,\Delta u + (\Delta u)^2$$

and, for $x = 0$,

$$\Delta u = (\Delta x)^2 \sin \frac{1}{\Delta x}.$$

If Δx is made to assume the sequence of values

$$\frac{1}{\pi}, \frac{1}{2\pi}, \frac{1}{3\pi}, \dots, \frac{1}{n\pi}, \dots,$$

then $\Delta u = 0$, and hence

$$\left(\frac{du}{dx}\right)_{x=0} = 0.$$

Moreover,

$$\frac{dz}{dx} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \frac{0}{\Delta x} = 0,$$

and the formula (16-2) is valid.

17. Derivatives and Differentials of Higher Orders. Inasmuch as the derivative of $y = f(x)$ is a function of x, it may be possible to differentiate again to produce a new function, which is called the *second derivative* of $f(x)$. The customary notations for the second derivative are $f''(x)$ and $\frac{d^2y}{dx^2}$. The third derivative, that is, the derivative of the second derivative, is denoted by $f'''(x)$ or $\frac{d^3y}{dx^3}$, and so on.

The reason for the notation $\frac{d^ny}{dx^n}$ will be seen from the definition of the differentials of higher order. The differential of $y = f(x)$,

$$dy = f'(x)\,dx,$$

is a function of x and dx, where dx is entirely arbitrary. Calculating the differential of $f'(x)$ and treating dx as a constant give the second differential

$$d^2y = d[f'(x)]\,dx = f''(x)\,(dx)^2.$$

Similarly,

$$d^3y = f'''(x)\,(dx)^3,$$

and so forth.

The formula for the differential of the nth order is

$$d^ny = f^{(n)}(x)\,(dx)^n, \tag{17-1}$$

so that one can write $f^{(n)}(x)$ as the quotient of two finite quantities, namely,

$$f^{(n)}(x) = \frac{d^n y}{dx^n}.$$

This notation, however, is bad. For consider a composite function $z = f(y)$ where $y = \varphi(x)$. From Sec. 16 it follows that

$$dz = f'(y)\,dy$$

when y is a dependent variable. But calculating the differentials of higher order by the formula for the differential of the product gives

$$\begin{aligned} d^2z &= d[f'(y)\,dy] = f'(y)\,d(dy) + dy\,d[f'(y)] \\ &= f'(y)\,d^2y + f''(y)\,(dy)^2, \end{aligned} \tag{17-2}$$

which differs from the formula (17-1).

If y is regarded as an independent variable, then dy must be considered as constant so that $d^2y = 0$ and (17-2) agrees with (17-1), but the two results are quite distinct if y is not an independent variable. This is one of the reasons why the differentials of higher orders are seldom used.*

PROBLEM

Let $y = f_1(t)$ and $x = f_2(t)$ be such that the functional dependence of y upon x is given in a parametric form. It is known that

$$\frac{dy}{dx} = \frac{f_1'(t)}{f_2'(t)}.$$

Now

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right).$$

Applying the rule for the differentiation of quotients gives

$$\frac{d^2y}{dx^2} = \frac{d^2y \cdot dx - d^2x \cdot dy}{(dx)^3} \tag{17-3}$$

On the other hand, using formula (17-1),

$$\begin{aligned} dy &= f_1'(t)\,dt, & d^2y &= f_1''(t)\,(dt)^2, \\ dx &= f_2'(t)\,dt, & d^2x &= f_2''(t)\,(dt)^2. \end{aligned}$$

---

* See also the problem on p. 48.

Substitute these expressions in the right-hand member of (17-3) and compare it with the expression for $\frac{d^2y}{dx^2}$ obtained in the usual way.

18. Fermat's Theorem. The reader is doubtless familiar with some geometrical considerations that lead one to suspect that a differentiable function $f(x)$ has a vanishing derivative at the points where $f(x)$ attains its maximum or minimum values. In fact, these geometrical considerations form a basis for the discussion of the maxima and minima in the first course in calculus. An analytical proof of the theorem due to Fermat gives a rigorous justification for such considerations.

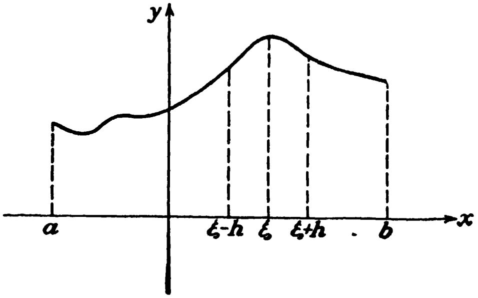

Fig. 12.

Theorem. *If a function $f(x)$ assumes a maximum or a minimum value at an interior point ξ of the interval (a, b), and if $f(x)$ is differentiable at $x = \xi$, then $f'(\xi) = 0$.*

Let it be supposed that at the point $x = \xi$ the function attains its maximum value. (The proof for the case of a minimum is entirely analogous and is obtained by reversing the signs of inequality.) Remembering that $x = \xi$ is an interior point of the interval and that $f(\xi)$ is a maximum, one has $f(\xi + h) - f(\xi) \leq 0$ for positive or negative h. Form the difference quotient

$$\frac{f(\xi + h) - f(\xi)}{h}. \tag{18-1}$$

Since the numerator of (18-1) is negative or zero,

$$\frac{f(\xi + h) - f(\xi)}{h} \leq 0, \qquad \text{for} \qquad h > 0, \tag{18-2}$$

$$\frac{f(\xi + h) - f(\xi)}{h} \geq 0 \qquad \text{for} \qquad h < 0. \tag{18-3}$$

Taking the limit of (18-2) as $h \to 0$ from the right, there results

$$f'(\xi) \leq 0,$$

while the limit of (18-3) as $h \to 0$ from the left is

$$f'(\xi) \geq 0.$$

But by hypothesis the derivative exists at $x = \xi$, so that the left-hand and right-hand limits must be equal, which is possible only if $f'(\xi) = 0$.

19. Rolle's Theorem.

Theorem. *If a function* $f(x)$ *is continuous in the interval* $a \leq x \leq b$ *and has a derivative at every interior point of* (a, b), *and if the values of the function are equal at the end points of the interval, then there exists at least one point* $x = \xi$, $(a < \xi < b)$, *such that* $f'(\xi) = 0$.

Since $f(x)$ is continuous in the closed interval (a, b), it attains its maximum value M and its minimum value m in (a, b) (Theorem 2, Sec. 13). If it happens that $m = M$, then it follows that

$$f(x) = \text{const.} = m.$$

But the derivative of the constant is zero so that the theorem is true in this case.

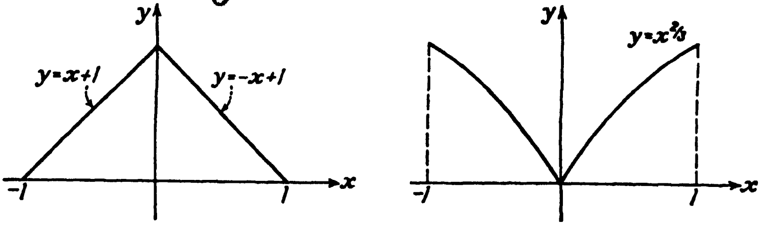

FIG. 13.

Next, assume that $m < M$. By hypothesis $f(a) = f(b)$ so that at least one of the values m or M is different from the value of the function at the end points. Let it be M. Then the maximum value M is attained, not at the end points of (a, b), but at some interior point ξ, and it follows from Fermat's theorem that $f'(\xi) = 0$.

It is important to observe that if the condition of the existence of the derivative is not fulfilled at a single interior point of (a, b), the theorem may not be valid. This can be seen from Fig. 13.

In neither of the examples in Fig. 13 does the derivative exist at $x = 0$.

If $f(a) = f(b) = 0$, Rolle's theorem can be stated as follows: *If a and b are two roots of the equation $f(x) = 0$, then $f'(x) = 0$ has at least one root between a and b, provided that $f(x)$ is continuous in the interval $a \leq x \leq b$ and has a derivative at every interior point of (a, b).*

✓20. Mean-value Theorem.

Theorem. *If a function $f(x)$ is continuous in the interval $a \leq x \leq b$ and possesses a derivative at every interior point of (a, b), then there exists a point ξ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \tag{20-1}$$

If $f(a) = f(b)$, the theorem is a special case of Rolle's theorem. Assume $f(a) \neq f(b)$, and construct a function

$$\underline{F(x) = kx + f(x)}$$

where the constant k will be so determined that $F(x)$ satisfies the conditions of Rolle's theorem, namely,

$$F(a) = F(b), \checkmark$$

or

$$ka + f(a) = kb + f(b).$$

Solving for k gives

$$k = -\frac{f(b) - f(a)}{b - a}.$$

With this choice of k, the function $F(x)$ satisfies the conditions of Rolle's theorem. Hence, there exists a point $x = \xi$ such that

$$F'(\xi) = k + f'(\xi) = 0.$$

Then

$$f'(\xi) = \frac{f(b) - f(a)}{b - a},$$

and the theorem is established. ✓

A geometrical interpretation of this theorem is interesting. The expression

$$\frac{f(b) - f(a)}{b - a}$$

is equal to the slope of the secant line joining the points A and B on the locus of $f(x)$ (Fig. 14).

The mean-value theorem asserts that there exists a point P on the locus of $f(x)$ such that the tangent line at this point is parallel to the secant line AB.

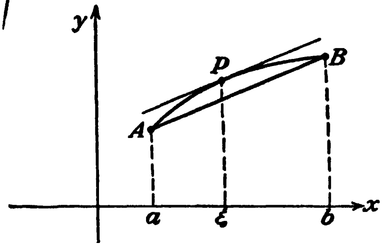

FIG. 14.

The mean-value theorem (20-1) can be written as

$$(20\text{-}2)\qquad f(b) - f(a) = f'(\xi)(b - a),$$

which gives the magnitude of the increment $f(b) - f(a)$. For this reason the form (20-2) of the mean-value theorem is frequently called the *formula of finite increments.*

Noting that ξ is an interior point of (a, b) and setting

$$\frac{\xi - a}{b - a} = \theta, \qquad \text{where} \qquad 0 < \theta < 1,$$

one can write

$$\xi = a + \theta(b - a).$$

Substituting this expression for ξ in (20-2) gives

$$(20\text{-}3)\qquad f(b) - f(a) = (b - a)f'[a + \theta(b - a)].$$

If the length of the interval (a, b) is denoted by h,

$$b = a + h$$

so that the expression (20-3) can be written in the following useful form:

$$(20\text{-}4)\qquad f(a + h) - f(a) = hf'(a + \theta h),$$

where $0 < \theta < 1$.

An important theorem follows immediately from the mean-value theorem.

Theorem. *If a function has a derivative that vanishes for all values of x in the interval $a \leq x \leq b$, the function is a constant.*

Moreover, if two functions have derivatives that are equal for all values of x in (a, b), then the functions differ by a constant.

Let x be any point of the interval (a, b), and assume that $f'(x) \equiv 0$ in (a, b). It follows from (20-2) that

$$f(x) - f(a) = f'(\xi)(x - a) = 0.$$

Hence,

$$f(x) = f(a),$$

which is constant.

If $f_1(x)$ and $f_2(x)$ are two functions with equal derivatives in (a, b), then

$$F(x) \equiv f_1(x) - f_2(x)$$

has a zero derivative in (a, b). Therefore,

$$F(x) = \text{const.},$$

and

$$f_1(x) = f_2(x) + \text{const.}$$

21. Theorem of Cauchy. L'Hospital's Rule.

Theorem. *If $f(x)$ and $\varphi(x)$ are continuous in the interval $a \leq x \leq b$ and have derivatives at each interior point of (a, b), then*

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)}, \qquad (a < \xi < b),$$

provided that

i. *$f'(x)$ and $\varphi'(x)$ do not vanish simultaneously in (a, b);*

ii. *$\varphi(b) \neq \varphi(a)$.*

Note that the function

$$F(x) \equiv f(x)[\varphi(b) - \varphi(a)] - \varphi(x)[f(b) - f(a)]$$

satisfies the conditions of Rolle's theorem, since $F(a) = F(b)$. Hence,

$$f'(\xi)[\varphi(b) - \varphi(a)] - \varphi'(\xi)[f(b) - f(a)] = 0, \tag{21-1}$$

where $(a < \xi < b)$.

Since $f'(x)$ and $\varphi'(x)$ cannot vanish for the same value of x, and since $\varphi(b) \neq \varphi(a)$, it follows that $\varphi'(\xi)$ cannot vanish. Consequently, one can divide (21-1) by $\varphi'(\xi)[\varphi(b) - \varphi(a)]$ to obtain the desired formula

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)}, \qquad (a < \xi < b). \tag{21-2}$$

If $\varphi'(x)$ is assumed to be different from zero at every interior point of the interval (a, b), then the conditions (i) and (ii) of the theorem are satisfied automatically. The condition (i) is obviously satisfied since $\varphi'(x) \neq 0$, and the fact that $\varphi(b) - \varphi(a) \neq 0$ follows directly from the mean-value theorem.

It is interesting to observe that (21-2) contains, as a special case, the mean-value theorem, for, setting $\varphi(x) = x$ gives

$$f(b) - f(a) = f'(\xi)(b - a).$$

It is readily shown that (21-2) can be written in the form

$$\frac{f(a + h) - f(a)}{\varphi(a + h) - \varphi(a)} = \frac{f'(a + \theta h)}{\varphi'(a + \theta h)}, \qquad (0 < \theta < 1).$$

An important use of formula (21-2) is in the evaluation of indeterminate forms. Thus, if $f(a) = \varphi(a) = 0$, and if the theorem be applied to the interval (a, x), there results

$$\frac{f(x)}{\varphi(x)} = \frac{f'(\xi)}{\varphi'(\xi)}, \qquad (a < \xi < x).$$

From this equation the following rule is obtained:

Rule. *If $f(a) = \varphi(a) = 0$, and if the ratio $\frac{f'(x)}{\varphi'(x)}$ approaches a limit as $x \to a$, then $\frac{f(x)}{\varphi(x)}$ approaches the same limit.* This rule for evaluating indeterminate forms is due to L'Hospital. By way of illustration consider the calculation of

$$\lim_{x \to 0} \frac{1 - \cos x}{x^2}.$$

Now $(1 - \cos x)_{x=0} = (x^2)_{x=0} = 0$, and one can apply the rule of L'Hospital provided that

$$\lim_{x \to 0} \frac{\sin x}{2x}$$

exists. But this again is an indeterminate form since

$$(\sin x)_{x=0} = (2x)_{x=0} = 0.$$

But

$$\lim_{x\to 0} \frac{\sin x}{2x}$$

exists since

$$\lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

Hence,

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}.$$

The rule of L'Hospital can also be applied to indeterminate forms of the type $\frac{\infty}{\infty}$. The proof is direct. Let $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} \varphi(x) = \infty$, and assume that $f(x)$ and $\varphi(x)$ have nonvanishing derivatives in the interval $a < x < x_0$.

Applying the formula of Cauchy to the interval (x, x_0) gives

$$\frac{f(x) - f(x_0)}{\varphi(x) - \varphi(x_0)} = \frac{f'(\xi)}{\varphi'(\xi)}, \qquad (x < \xi < x_0).$$

But

$$\frac{f(x) - f(x_0)}{\varphi(x) - \varphi(x_0)} = \frac{f(x)}{\varphi(x)} \cdot \frac{1 - \dfrac{f(x_0)}{f(x)}}{1 - \dfrac{\varphi(x_0)}{\varphi(x)}}.$$

Thus

$$\text{(21-3)} \qquad \frac{f(x)}{\varphi(x)} = \frac{1 - \dfrac{\varphi(x_0)}{\varphi(x)}}{1 - \dfrac{f(x_0)}{f(x)}} \cdot \frac{f'(\xi)}{\varphi'(\xi)}.$$

Assume that

$$\lim_{x\to a} \frac{f'(x)}{\varphi'(x)} = L.$$

This means that one can choose x_0 so near a that for all values of x between a and x_0

$$\left|\frac{f'(x)}{\varphi'(x)} - L\right| < \epsilon.$$

Having chosen x_0 to satisfy this inequality, let $x \rightarrow a$; then the first factor in the right-hand member of (21-3) can be made as near unity as desired. Hence, the left-hand member of (21-3) can be made to differ from L by as little as desired. This is another way of saying that

$$\lim_{x\rightarrow a} \frac{f(x)}{\varphi(x)} = L.$$

The foregoing proof is easily modified to apply when $a = \infty$.

Example 1.

$$\lim_{x\rightarrow 0+} x \log x = \lim_{x\rightarrow 0+} \frac{\log x}{\frac{1}{x}} = \lim_{x\rightarrow 0+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0.$$

Example 2. Evaluate $\lim_{x\rightarrow \infty} \frac{x - \sin x}{x}$.

Applying the rule of L'Hospital gives

$$\lim_{x\rightarrow \infty} \frac{x - \sin x}{x} = \lim_{x\rightarrow \infty} \frac{1 - \cos x}{1},$$

which does not approach a limit since $\cos x$ oscillates as $x \rightarrow \infty$. But

$$\lim_{x\rightarrow \infty} \frac{x - \sin x}{x} = \lim_{x\rightarrow \infty} \left(1 - \frac{\sin x}{x}\right) = 1$$

since

$$\left|\frac{\sin x}{x}\right| < \frac{1}{|x|} \rightarrow 0, \qquad \text{as} \qquad x \rightarrow \infty.$$

Note that this does not constitute a contradiction since the theorem asserts that *if the quotient of the derivatives approaches a limit, then the quotient of the functions has the same limit.*

PROBLEMS

1. Show that

(*a*) $\lim_{x\rightarrow \infty} \frac{\log x}{x^n} = 0$, if $n > 0$;

(*b*) $\lim_{x\rightarrow 0} \frac{x^n}{e^x - 1} = 0$, if $n > 1$;

(*c*) $\lim\limits_{x \to \infty} x^n e^{-x} = 0$, if $n > 0$;

(*d*) $\lim\limits_{x \to 0+} x^x = 1$. *Hint:* Evaluate $\lim\limits_{x \to 0+} \log x^x$;

(*e*) $\lim\limits_{x \to 0} \left(\dfrac{1}{x^2} - \dfrac{\cot x}{x}\right) = \dfrac{1}{3}$.

2. Find

(*a*) $\lim\limits_{x \to \infty} x^n \log x$;

(*b*) $\lim\limits_{x \to \infty} x^n e^{-x}$;

(*c*) $\lim\limits_{x \to 0} \dfrac{\sin x}{x^n}$;

(*d*) $\lim\limits_{x \to \infty} \dfrac{x^2}{x - \sin x}$;

(*e*) $\lim\limits_{x \to 0} \left(\dfrac{1}{x} - \cot x\right)$.

3. Show that

(*a*) $\lim\limits_{x \to 0} \left(\dfrac{\tan x}{x}\right)^{x^{-2}} = \sqrt[3]{e}$;

(*b*) $\lim\limits_{x \to \frac{\pi}{2}} (\tan x)^{\sin 2x} = 1$

CHAPTER III

FUNCTIONS OF SEVERAL VARIABLES

22. Limits and Continuity. Let $u = f(x, y)$ be a real, single-valued function of two independent variables x and y. One can think of the pair of values $x = x_0$ and $y = y_0$ as representing a point in the xy-plane, so that $u = f(x, y)$ at (x_0, y_0) becomes a function of the point.

The function $f(x, y)$ may be determined for every point of the xy-plane, or the totality of admissible points may occupy a certain region R in the xy-plane. Thus, if the attention is confined to real values only, the function

$$u = \sqrt{1 - x^2 - y^2}$$

is defined for those values of x and y for which

$$x^2 + y^2 \leq 1.$$

The region R in this case is a circle of radius unity whose center is at the origin of the xy-plane. If the totality of points making up the region R includes the boundary of the region, then R is said to be closed. The function

$$u = \frac{1}{\sqrt{1 - x^2 - y^2}}$$

is defined in the open region $x^2 + y^2 < 1$.

The definition of continuity of a function of two independent variables is a generalization of the definition of continuity of a function of a single variable. Let $f(x, y)$ be defined at the point (x_0, y_0) and in a sufficiently small region about this point. If

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = f(x_0, y_0), \tag{22-1}$$

the function $f(x, y)$ is said to be continuous at the point (x_0, y_0). The precise arithmetical meaning of the statement embodied in (22-1) is the following:

If for any preassigned positive number ϵ *one can find a positive number* δ *such that*

$$|f(x, y) - f(x_0, y_0)| < \epsilon,$$

whenever $|x - x_0| < \delta$ *and* $|y - y_0| < \delta$, *then the function is continuous.*

Geometrically this means (Fig. 15) that one can find a square about the point (x_0, y_0) so small that the difference between the values of the function at (x_0, y_0) and at any other point (x, y) within the square will be in absolute value less than any preassigned number ϵ.

Inasmuch as nothing is said in the definition (22-1) about the manner in which $x \rightarrow x_0$ and $y \rightarrow y_0$, the limit must exist for any mode of approach to (x_0, y_0).

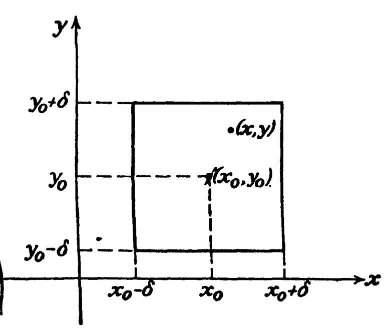

FIG. 15.

A function that is continuous at every point of a region R is said to be *continuous in the region R*. If the region is closed, the function is continuous in the closed region. For example, $u = \sqrt{1 - x^2 - y^2}$ is continuous within and upon the boundary of the region bounded by the circle $x^2 + y^2 = 1$.

Only real functions of real variables x and y are considered in this chapter. There are theorems regarding such functions that are entirely analogous to those given in Chap. I for functions of a single variable. The more important of these are the following:*

Theorem 1. *If* $f(x, y)$ *is continuous at* (x_0, y_0) *and* $f(x_0, y_0) \neq 0$, *then one can find a region about the point* (x_0, y_0) *such that* $f(x, y)$ *has the same sign as* $f(x_0, y_0)$ *throughout the region.*

Theorem 2. *A function* $f(x, y)$ *that is continuous in a closed region is bounded in that region.*

Theorem 3. *If* $f(x, y)$ *is continuous in a closed region, then there exists in this region at least one point* (x, y) *at which* $f(x, y)$

---

* The reasoning used in establishing the validity of these theorems is of the same character as that employed in proving the corresponding theorems for functions of a single variable.

takes its maximum value and at least one point where it assumes its minimum value.

Theorem 4. *If u and v are continuous functions of x and y, and if z is a continuous function of u and v, then z is a continuous function of x and y.*

The definition of uniform continuity of a function of two variables is as follows:

If for any preassigned positive number ϵ, one can find a positive number δ, the same for the entire region R, such that

$$|f(x_1, y_1) - f(x_2, y_2)| < \epsilon,$$

for every pair of values (x_1, y_1) and (x_2, y_2) that satisfy the inequalities $|x_1 - x_2| < \delta$ and $|y_1 - y_2| < \delta$, then $f(x, y)$ is uniformly continuous in R.

A function of two variables that is continuous in a closed region R can be shown to be uniformly continuous in R.

The extension of these definitions to functions of more than two variables is immediately obvious. If

$$u = f(x, y, z),$$

the triplets of values (x, y, z) can be associated with the points of three-dimensional space. The square region $|x - x_0| < \delta$ and $|y - y_0| < \delta$ used in the definition of continuity of $f(x, y)$ must be replaced by a cubical region defined by the inequalities

$$|x - x_0| < \delta, \qquad |y - y_0| < \delta, \qquad |z - z_0| < \delta.$$

If the function u depends on more than three independent variables, the usual geometrical interpretation of the independent variables as the coordinates of points in ordinary space fails, but the geometrical language may still be used. Thus $u = f(x_1, x_2, x_3, x_4)$, where x_1, x_2, x_3, and x_4 are four independent variables, may be thought of as defined over a four-dimensional manifold, and the quadruplets of numbers (x_1, x_2, x_3, x_4) may be associated with the points of a hyperspace.

A continuous function $u = f(x, y)$ of two independent variables x and y is a continuous function of each of the variables taken separately. Thus, if the value of one of the variables is fixed by setting $y = y_0$, the resulting function of the single variable x, namely,

$$u = f(x, y_0),$$

is obviously continuous in x. The converse of this statement may not be true, however. For example, let

$$\begin{aligned} f(x, y) &= \frac{2xy}{x^2 + y^2}, \text{ unless } x = 0 \text{ and } y = 0 \text{ simultaneously,} \\ &= 0, \text{ if } x = 0 \text{ and } y = 0. \end{aligned}$$

Then for any fixed value of y (or x) the resulting function of a single variable is easily shown to be continuous. However, $f(x, y)$ is not a continuous function of the two variables x and y at $x = 0$ and $y = 0$. To show this, it will suffice to demonstrate that there exists one mode of approach toward the origin for which $\lim f(x, y) \neq 0$. Let the point (x, y) approach $(0, 0)$ along the line $y = x$. Then

$$\lim_{\substack{x \to 0 \\ y \to 0}} \frac{2xy}{x^2 + y^2} = \lim_{x \to 0} \frac{2x^2}{2x^2} = 1,$$

whereas $f(x, y)$ is defined to be zero at the origin.

A word of caution is in order in regard to the meaning of the symbols:

(a) $$\lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y),$$

(b) $$\lim_{x \to x_0} \lim_{y \to y_0} f(x, y),$$

and

(c) $$\lim_{y \to y_0} \lim_{x \to x_0} f(x, y).$$

The limits in (b) and (c) are called the *repeated* or *iterated limits*, and the order of the two separate limiting processes in (b) and (c) cannot be interchanged in general. Thus, consider

$$f(x, y) = \frac{x + y}{x - y}.$$

Then

$$\lim_{y \to 0} \left(\lim_{x \to 0} \frac{x + y}{x - y} \right) = -1,$$

while

$$\lim_{x \to 0} \left(\lim_{y \to 0} \frac{x + y}{x - y} \right) = 1.$$

The limit in (*a*), on the other hand, must exist for any conceivable mode of approach of (x, y) to (x_0, y_0). Hence, if the limit in (*a*) exists, the corresponding limits in (*b*) and (*c*) will exist and will be equal if the inner limits exist. The converse, however, is not true. The repeated limits (*b*) and (*c*) may exist and be equal, and yet the limit in (*a*) may fail to exist.*

PROBLEM

Show that $f(x, y) = \dfrac{x^2y^2}{x^2y^2 + (x - y)^2}$ is a continuous function in each of the variables taken separately but is not a continuous function of both variables at $(0, 0)$.

23. Partial Derivatives. Let $u = f(x, y)$ be a single-valued function of the independent variables x and y, and let it be defined at some point (x_0, y_0) and for all values of (x, y) in some region† R about the point (x_0, y_0). If y is set equal to y_0, u becomes a function of the single variable x, namely,

$$u = f(x, y_0).$$

If this function has a derivative with respect to the variable x, the derivative is called the *partial derivative of $f(x, y)$ with respect to x, for $y = y_0$*. In like manner, if x is assigned a constant value x_0, the derivative with respect to y of the resulting function $f(x_0, y)$ is called the *partial aerivative of $f(x, y)$ with respect to y, for $x = x_0$*. The customary notations for the partial derivative of $u = f(x, y)$ with respect to x are:

$$\frac{\partial u}{\partial x},\ u_x,\ f_x(x, y), \text{ and } \frac{\partial f}{\partial x}.$$

The partial derivatives of a function $f(x_1, x_2, \cdots, x_n)$, of n independent variables, are obtained by fixing the values of $n - 1$ of the variables and calculating the derivative of the resulting function of a single variable. Thus,

$$f(x, y, z) = yx^2 - 2yx + xz^2$$

---

* See in this connection Hobson, E. W., Theory of Functions of a Real Variable, p. 408.

† Analogous to the definition of the "neighborhood" or "vicinity" of a point given in Sec. 3 for a function of a single variable, the totality of all points whose distance from a given point P is less than a given number $\epsilon > 0$ is called *a neighborhood of the point P*.

has the partial derivatives

$$\frac{\partial f}{\partial x} = 2xy - 2y + z^2,$$
$$\frac{\partial f}{\partial y} = x^2 - 2x,$$
$$\frac{\partial f}{\partial z} = 2xz.$$

Recalling the definition of the derivative of a function of a single variable, one can see that the expressions

$$\frac{\partial u}{\partial x} = \lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h},$$

and

$$\frac{\partial u}{\partial y} = \lim_{h\to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h},$$

define the partial derivatives of $f(x, y)$ at the point (x_0, y_0). Even though the notation $\frac{\partial u}{\partial x}$ suggests a fraction, it must be noted carefully that it is merely a symbol for the partial derivative, which does not find a sensible interpretation as a quotient analogous to that given for the ordinary derivative.

Let Δu denote the change in the value of $u = f(x, y)$ when x and y acquire, respectively, the increments Δx and Δy. Then

$$\Delta u = f(x + \Delta x, y + \Delta y) - f(x, y). \tag{23-1}$$

Adding and subtracting $f(x, y + \Delta y)$ in the right-hand member of (23-1) give

$$\Delta u = [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] + [f(x, y + \Delta y) - f(x, y)].$$

The expression in the first bracket represents the increment received by the function $f(x, y + \Delta y)$ when x undergoes a change of magnitude Δx; whereas the second bracket gives the increment of the function $f(x, y)$ when x is kept fixed.

The application of the mean-value theorem gives

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = f_x(x + \theta_1 \Delta x, y + \Delta y)\,\Delta x,$$
$$f(x, y + \Delta y) - f(x, y) = f_y(x, y + \theta_2 \Delta y)\,\Delta y,$$

where θ_1 and θ_2 are numbers between 0 and 1. Then Δu can be written in the form

$$\Delta u = f_x(x + \theta_1 \Delta x, y + \Delta y)\,\Delta x + f_y(x, y + \theta_2 \Delta y)\,\Delta y. \tag{23-2}$$

If it is assumed that the partial derivatives f_x and f_y are continuous functions of x and y, then

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} f_x(x + \theta_1 \Delta x, y + \Delta y) = f_x(x, y),$$

and

$$\lim_{\Delta y \to 0} f_y(x, y + \theta_2 \Delta y) = f_y(x, y).$$

The last two equations can be rewritten to read

$$\begin{cases} f_x(x + \theta_1 \Delta x, y + \Delta y) = f_x(x, y) + \eta_1 \\ f_y(x, y + \theta_2 \Delta y) \qquad\quad = f_y(x, y) + \eta_2, \end{cases} \tag{23-3}$$

where η_1 and η_2 tend to zero when $\Delta x \to 0$ and $\Delta y \to 0$.

Substituting from (23-3) in (23-2) gives

$$\Delta u = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \eta_1 \Delta x + \eta_2 \Delta y. \tag{23-4}$$

The sum of the first two terms in the right-hand member of (23-4) is called the *principal part of the increment* Δu, or the *total differential of* u, and is denoted by the symbol

$$du = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \tag{23-5}$$

As in the case of a function of a single variable, the increments and the differentials of the independent variables are identical, so that $dx \equiv \Delta x$ and $dy \equiv \Delta y$.

Since η_1 and η_2 are functions of Δx and Δy that tend to zero with Δx and Δy, the expressions

$$\eta_1 \Delta x \text{ and } \eta_2 \Delta y$$

are infinitesimals of higher order* than Δx and Δy. Accordingly, for small values of Δx and Δy, the value of Δu is very nearly that of du.†

---

* It will be recalled that a variable whose limit is zero is called an *infinitesimal.* If $\lim \frac{\alpha}{\beta} = 0$, where α and β are infinitesimals, α is said to be of higher order than β.

† See Prob. 6, p. 66.

Formula (23-4) is analogous to that developed in Sec. 15 for the increment of the function of a single variable x, where it was found that

$$\Delta y = f'(x)\,\Delta x + \eta\,\Delta x.$$

Since $u = f(x, y)$, (23-5) can be written as

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy. \tag{23-6}$$

The extension of formula (23-6) to functions of any number of variables is immediate. If $u = f(x_1, x_2, \cdots, x_n)$, where $x_1, x_2, \cdots, x_n$, are the independent variables, the differential of u is defined as

$$du = \frac{\partial u}{\partial x_1}\,dx_1 + \frac{\partial u}{\partial x_2}\,dx_2 + \cdots + \frac{\partial u}{\partial x_n}\,dx_n. \tag{23-7}$$

Theorem. *If* $x_1, x_2, \cdots, x_n$ *are independent variables and* $u = f(x_1, x_2, \cdots, x_n)$ *is continuous together with its partial derivatives, and if*

$$\begin{aligned} du = {}& P_1(x_1, x_2, \cdots, x_n)\,dx_1 + P_2(x_1, x_2, \cdots, x_n)\,dx_2 \\ &+ \cdots + P_n(x_1, x_2, \cdots x_n)\,dx_n, \end{aligned} \tag{23-8}$$

then

$$P_1 = \frac{\partial u}{\partial x_1}, \qquad P_2 = \frac{\partial u}{\partial x_2}, \cdots, P_n = \frac{\partial u}{\partial x_n}.$$

Since the variables are independent, one can set

$$dx_2 = dx_3 = \cdots = dx_n = 0$$

in (23-7) and (23-8) to obtain

$$du = \frac{\partial u}{\partial x_1}\,dx_1 = P_1\,dx_1$$

Thus

$$P_1 = \frac{\partial u}{\partial x_1}, \text{ and so forth.}$$

Corollary. *If* $du = 0$, *then* $\dfrac{\partial u}{\partial x_1} = \dfrac{\partial u}{\partial x_2} = \cdots = \dfrac{\partial u}{\partial x_n} \equiv 0.$

PROBLEMS

1. Find the partial derivatives of the following functions:

(a) $\log (x^2 + y^2 + z^2)$;
(b) $e^z \cos xy$;
(c) $x^3 - 3x^2y + z^3$;

(d) $\frac{e^x}{z^2 - y^2}$;

(e) $\frac{e^{xy}}{e^x + e^y}$;

(f) $\sin^{-1} \frac{y}{x}$;

(g) $\tan^{-1} \frac{x}{y}$;

(h) $\frac{x - y}{x + y}$;

(i) $e^{-x} \sin y$.

2. Find the differentials of the following functions:

(a) $(x^2 + y^2 + z^2)^{1/2}$;
(b) $(x^2 + y^2 + z^2)^{-1/2}$;
(c) $x \log (x^2 + z^2)$;
(d) $\log \tan (x^2 + y^2)$;
(e) $\frac{xy}{x + y}$.

3. What are the values of the partial derivatives of

$$z = \frac{x - y + 1}{x + y - 1}$$

at (0, 0)?

4. Find the partial derivatives of

$$\begin{aligned} x &= r \cos \theta \sin \varphi, \\ y &= r \sin \theta \sin \varphi, \\ z &= r \cos \theta, \end{aligned}$$

with respect to r, θ and φ.

5. If $u = x^2y + y^2z + z^2x$, verify that

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = (x + y + z)^2.$$

6. It was remarked in Sec. 23 that for small changes in the values of the independent variables the change in the value of the function is nearly equal to the magnitude of the differential. This fact can be used to calculate the approximate error in the quantity $f(x, y, z)$ determined by measurements of x, y, and z. The error in f, due to small errors dx, dy, and dz in x, y, and z, is approximately equal to

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$

Show that the approximate relative error $\frac{df}{f}$ is equal to $d \log f$.

7. Referring to Prob. 6, find the approximate error in calculating the area of a triangular piece of land, two sides of which are measured as 125 and 100 ft., respectively, and the included angle as 60.° The possible error in measuring the sides is 0.2 ft. and that in the measure of the angle is 1°. *Hint:* $A = \frac{1}{2}xy \sin \alpha$. *Ans.* 74.0 sq. ft.

8. In estimating the cost of a pile of bricks measured as 6 × 50 × 4 ft., the tape is stretched 1 per cent beyond the estimated length. If the count is 12 bricks to 1 cu. ft. and bricks cost $8 per thousand, find the error in cost. *Ans.* $3.46.

9. In determining specific gravity by the formula $s = \frac{A}{A - W}$, where A is the weight in air and W is the weight in water, A can be read within 0.01 lb. and W within 0.02 lb. Find approximately the maximum error in s if the readings are $A = 1.1$ lb. and $W = 0.6$ lb. Find the maximum relative error $\frac{\Delta s}{s}$. *Ans.* 0.112; 0.054.

10. The period of a simple pendulum with small oscillations is

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

If T is computed using $l = 8$ ft. and $g = 32$ ft. per second per second, find the approximate error in T if the true values are $l = 8.05$ ft. and $g = 32.01$ ft. per second per second. Find also the percentage error. *Ans.* 0.003π; 0.3 per cent.

11. The diameter and altitude of a can in the shape of a right circular cylinder are measured as 4 in. and 6 in., respectively. The probable error in each measurement is 0.1 in. Find approximately the maximum possible error in the values computed for the volume and the lateral surface. *Ans.* 1.6π; π.

12. Show that the relative error of the product is equal to the sum of the relative errors of the factors.

24. Differentiation of Composite Functions. Let $u = f(x, y)$ be a function of the variables x and y which in turn are functions of some independent variable t. If t is given an increment Δt, the functions x and y will acquire increments Δx and Δy, and, consequently, u will receive an increment Δu.

Assuming that $u = f(x, y)$ is continuous together with its partial derivatives, one can write [see (23-4)]

$$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \eta_1 \Delta x + \eta_2 \Delta y.$$

Dividing both sides of this expression by Δt gives

$$\frac{\Delta u}{\Delta t} = \frac{\partial u}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial u}{\partial y}\frac{\Delta y}{\Delta t} + \eta_1\frac{\Delta x}{\Delta t} + \eta_2\frac{\Delta y}{\Delta t}. \tag{24-1}$$

Now if it be supposed that x and y can be differentiated with respect to t, the expression (24-1) gives, upon passing to the limit as $\Delta t \rightarrow 0$,

$$(24\text{-}2) \qquad \begin{cases} \dfrac{du}{dt} = \dfrac{\partial u}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial u}{\partial y}\dfrac{dy}{dt} \\ \quad \equiv \dfrac{\partial f}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial f}{\partial y}\dfrac{dy}{dt}, \end{cases}$$

since $\eta_1 \rightarrow 0$ and $\eta_2 \rightarrow 0$. The reason for the vanishing of η_1 and η_2 as $\Delta t \rightarrow 0$ is that η_1 and η_2 are functions of Δx and Δy, and since x and y are assumed to be differentiable functions of t, they are surely continuous in t.

Formula (24-2) gives the rule for the differentiation of composite functions. It is clear that if u is a function of a set of variables, $x_1, x_2, \cdots, x_n$, where each variable is a function of an independent variable t, the derivative of u with respect to t is given by the formula

$$\frac{du}{dt} = \frac{\partial u}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial u}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial u}{\partial x_n}\frac{dx_n}{dt}.$$

A special case of formula (24–2) is of interest. If it is assumed that $t = x$, (24-2) becomes

$$(24\text{-}3) \qquad \frac{du}{dx} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}\frac{dy}{dx}.$$

Formula (24-3) can be used to calculate the derivative of the implicit function given by

$$(24\text{-}4) \qquad f(x, y) = 0.$$

Let it be assumed* that (24-4) can be solved for y to yield a real solution

$$(24\text{-}5) \qquad y = \varphi(x);$$

then the substitution of (24-5) in the left-hand member of (24-4) gives an identity $0 = 0$, or

$$f[x, \varphi(x)] \equiv 0.$$

Thus the constant zero defines an implicit function of x, namely,

---

* For general theorems regarding the conditions to be satisfied by $f(x, y)$ if solution is to be possible, see Chap. XII.

(24-6) $$0 = f(x, y), \qquad \text{where} \qquad y = \varphi(x).$$

Applying (24-3) to (24-6) gives

$$0 = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx},$$

and solving for $\frac{dy}{dx}$,

(24-7) $$\frac{dy}{dx} = -\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}.$$

The formula (24-7) implies that $\frac{\partial f}{\partial y}$ must not vanish for the point (x_0, y_0) at which the derivative is calculated.

Example 1. Let $f(x, y) = 3x^3y^2 + x \cos y = 0$.

$$\frac{\partial f}{\partial x} = 9x^2y^2 + \cos y, \qquad \frac{\partial f}{\partial y} = 6x^3y - x \sin y,$$

so that

$$\frac{dy}{dx} = -\frac{9x^2y^2 + \cos y}{6x^3y - x \sin y},$$

for all values of x and y that satisfy the equation

$$3x^3y^2 + x \cos y = 0,$$

and for which $6x^3y - x \sin y \neq 0$.

Example 2. Let $x^2 + y^2 = 0$; then $\frac{\partial f}{\partial x} = 2x$, $\frac{\partial f}{\partial y} = 2y$. But it does not follow that

$$\frac{dy}{dx} = -\frac{x}{y}.$$

This result is absurd inasmuch as the only real values of x and y that satisfy $x^2 + y^2 = 0$ are $x = 0$ and $y = 0$. Since $\frac{\partial f}{\partial y}$ vanishes at this point, the formal procedure used in obtaining $\frac{dy}{dx}$ is meaningless.

Example 3. Let $f(x, y) = 0$ represent the locus of some curve and let $P(x_0, y_0)$ be some point on the curve. The equation of the tangent line to the curve at the point P is

$$y - y_0 = \left(\frac{dy}{dx}\right)_{x=x_0}(x - x_0).$$

It follows from (24-7) that this equation can be written in the form

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

PROBLEMS

1. Find the equation of the tangent line to the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

at the point (x_0, y_0).

2. Find the equation of the tangent line to the folium of Descartes

$$x^3 + y^3 - 3axy = 0.$$

Note particularly the behavior of the tangent line to the folium at (0, 0).

3. Find $\frac{du}{dt}$, if

$$u = \tan^{-1}\frac{y}{x} \quad \text{and} \quad \begin{cases} y = e^t + e^{-t} \\ x = e^t - e^{-t}. \end{cases}$$

4. Find the equation of the tangent line to the ellipse

$$\begin{aligned} x &= a\cos\theta, \\ y &= b\sin\theta, \end{aligned}$$

at the point where $\theta = \frac{\pi}{4}$.

5. (*a*) Find $\frac{du}{dt}$, if $u = e^x \sin yz$ and $x = t^2$, $y = t - 1$, $z = \frac{1}{t}$;

(*b*) find $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial \theta}$, if $u = x^2 - 4y^2$, $x = r\sec\theta$, and $y = r\tan\theta$.

6. (*a*) Find $\frac{\partial u}{\partial x}$ and $\frac{du}{dx}$, if $u = x^2 + y^2$ and $y = \tan x$;

(*b*) given $V = f(x, y, z)$, where $x = r\cos\theta$, $y = r\sin\theta$, $z = t$; compute $\frac{\partial V}{\partial r}$, $\frac{\partial V}{\partial \theta}$, $\frac{\partial V}{\partial t}$ in terms of $\frac{\partial V}{\partial x}$, $\frac{\partial V}{\partial y}$, and $\frac{\partial V}{\partial z}$.

7. The equation of a perfect gas is $pv = RT$. At a certain instant a given amount of gas has a volume of 16 cu. ft. and is under a pressure of 36 lb. per square inch. Assuming $R = 10.71$, find the temperature T.

If the volume is increasing at the rate of 1/3 cu. ft. per second, and the pressure is decreasing at the rate 1/8 lb. per square inch per second, find the rate at which the temperature is changing. *Ans.* 53.78; 0.93.

25. Differentiation of Composite and Implicit Functions. The reasoning employed in the preceding section can be applied in obtaining the total differential, and hence the derivative, of a function of n variables

$$u = f(x_1, x_2, \cdots, x_n),$$

where

$$x_i = x_i(t), \qquad (i = 1, 2, \cdots, n),$$

are n differentiable functions of a single variable t.

The resulting expression for the total differential is

$$du = \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 + \cdots + \frac{\partial f}{\partial x_n}\,dx_n. \tag{25-1}$$

A question arises concerning the validity of formula (25-1) in the case where the variables x_i are functions of several independent variables $t_1, t_2, \cdots, t_m$. Thus, let

$$u = f(x_1, x_2, \cdots, x_n) \tag{25-2}$$

be a function of the n variables x_i, where the x_i are functions of the variables $t_1, t_2, \cdots, t_m$, say

$$x_i = x_i(t_1, t_2, \cdots, t_m), \qquad (i = 1, 2, \cdots, n). \tag{25-3}$$

If all of the variables except one, say t_k, are held fast, (25-2) becomes a function of the single variable t_k and one can calculate the derivative $\frac{\partial f}{\partial t_k}$ with the aid of (25-1). The notation $\frac{\partial f}{\partial t_k}$, instead of $\frac{df}{dt_k}$, is used to signify the fact that all variables except t_k are held fast.

Assuming the continuity of the derivatives involved, one can write

$$\begin{aligned}
\frac{\partial f}{\partial t_1} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_1},\\
\frac{\partial f}{\partial t_2} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_2},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\frac{\partial f}{\partial t_m} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_m} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_m} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_m}.
\end{aligned}$$

If $\frac{\partial f}{\partial t_1}, \frac{\partial f}{\partial t_2}, \cdots, \frac{\partial f}{\partial t_m}$ be multiplied respectively by dt_1, dt_2, $\cdots$, dt_m and the resulting expressions added, one obtains

$$\begin{aligned}\frac{\partial f}{\partial t_1}dt_1 &+ \frac{\partial f}{\partial t_2}dt_2 + \cdots + \frac{\partial f}{\partial t_m}dt_m \\ &= \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial t_1}dt_1 + \frac{\partial x_1}{\partial t_2}dt_2 + \cdots + \frac{\partial x_1}{\partial t_m}dt_m\right) \\ &+ \frac{\partial f}{\partial x_2}\left(\frac{\partial x_2}{\partial t_1}dt_1 + \frac{\partial x_2}{\partial t_2}dt_2 + \cdots + \frac{\partial x_2}{\partial t_m}dt_m\right) \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+ \frac{\partial f}{\partial x_n}\left(\frac{\partial x_n}{\partial t_1}dt_1 + \frac{\partial x_n}{\partial t_2}dt_2 + \cdots + \frac{\partial x_n}{\partial t_m}dt_m\right).\end{aligned}$$

The left-hand member of this expression is the total differential of $f(x_1, x_2, \cdots, x_n)$, regarded as a function of the independent variables $t_1, t_2, \cdots, t_m$, whereas the terms in the parentheses in the right-hand member are precisely the total differentials of (25-3). Hence, one can write

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n,$$

which shows that the formula (25-1) is valid whether the x_i's are the independent variables or are functions of any set of other independent variables.

The foregoing can be summarized as follows:

Theorem. *If* $u = f(x_1, x_2, \cdots, x_n)$, *then*

$$du = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n,$$

regardless of whether the variables x_i *are the independent variables or are functions of other independent variables* t_k. *It is understood that all the derivatives involved* $\left(\text{the } \frac{\partial f}{\partial x_k} \text{ and } \frac{\partial x_i}{\partial t_k}\right)$ *are continuous functions.*

Example. If $f(x, y) = x^2 + y^2$, where $x = r\cos\varphi$ and $y = r\sin\varphi$, then

$$\begin{aligned}\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} &= 2x\cos\varphi + 2y\sin\varphi \\ &= 2r\cos^2\varphi + 2r\sin^2\varphi = 2r,\end{aligned}$$

$$\frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \varphi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \varphi} = 2x(-r \sin \varphi) + 2y(r \cos \varphi)$$
$$= -2r^2 \cos \varphi \sin \varphi + 2r^2 \cos \varphi \sin \varphi = 0.$$

Also

$$df = 2r\,dr \qquad \text{or} \qquad df = 2x\,dx + 2y\,dy.$$

Let $f(x, y, z) = 0$ define any one of the variables as an implicit function of the remaining ones. If x and y are thought to be the independent variables and one can obtain* (at least theoretically) a real solution for z in terms of x and y, it is possible to write

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy.$$

But

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz = 0.$$

Substituting the value of dz in this equation gives

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy\right) = 0,$$

or

$$\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y}\right)dy = 0.$$

By the corollary to the theorem of Sec. 23, since x and y are independent variables,

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0$$

and

$$\frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0.$$

If $\frac{\partial f}{\partial z} \neq 0$, these equations give

$$\text{(25-4)} \qquad \frac{\partial z}{\partial x} = -\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial z}}, \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial f}{\partial y}}{\dfrac{\partial f}{\partial z}}.$$

The formulas (25-4) permit one to calculate the partial derivatives of the function z defined implicitly by an equation

$$f(x, y, z) = 0.$$

---

* See Sec. 113.

As an illustration, let

$$x^2 + 2y^2 - 3xz + 1 = 0.$$

Then, by (25-4),

$$\frac{\partial z}{\partial x} = -\frac{2x - 3z}{-3x}, \qquad \text{and} \qquad \frac{\partial z}{\partial y} = -\frac{4y}{-3x}.$$

PROBLEMS

1. Find $\dfrac{\partial u}{\partial r}$ and $\dfrac{\partial u}{\partial \theta}$, if $u = x^2 - 4y^2$, $x = r \sec \theta$, and $y = r \tan \theta$.

2. If $V = f(x, y, z)$, where $x = r \cos \theta$, $y = r \sin \theta$, $z = t$, find

$$\frac{\partial V}{\partial r}, \frac{\partial V}{\partial \theta}, \frac{\partial V}{\partial t}$$

in terms of $\dfrac{\partial V}{\partial x}$, $\dfrac{\partial V}{\partial y}$, and $\dfrac{\partial V}{\partial z}$.

3. (*a*) Find $\dfrac{dy}{dx}$, if $x \sec y + x^3 y = 1$;

(*b*) find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$, if

$$x^3 y - \sin z + z^2 = 0.$$

4. If f is a function of u and v, where $u = \sqrt{x^2 + y^2}$ and $v = \tan^{-1} \dfrac{y}{x}$, find $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, $\sqrt{\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2}$

5. If f is a function of u and v, where $u = r \cos \theta$ and $v = r \sin \theta$, find

$$\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}, \sqrt{\left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \theta}\right)^2}.$$

6. If $x = x' \cos \theta - y' \sin \theta$, $y = x' \sin \theta + y' \cos \theta$, prove that

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = \left(\frac{\partial f}{\partial x'}\right)^2 + \left(\frac{\partial f}{\partial y'}\right)^2.$$

7. Find the total differential, if $u = x^2 + y^2$, $x = r \cos \theta$, and $y = r \sin \theta$.

8. If $f = e^{xy}$, where $x = \log (u^2 + v^2)$ and $y = \tan^{-1} \dfrac{u}{v}$, find $\dfrac{\partial f}{\partial u}$ and $\dfrac{\partial f}{\partial v}$.

9. If $z = \dfrac{u + v}{1 - uv}$, $u = y \sin x$, and $v = e^{vx}$, find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$.

10. Find $\dfrac{\partial z}{\partial r}$ and $\dfrac{\partial z}{\partial s}$, if $z = \dfrac{x - y}{1 + xy}$, $x = \tan (r - s)$, and $y = e^{rs}$.

26. Euler's Theorem. A function f of n variables $x_1, x_2, \cdots, x_n$ is called *homogeneous of degree m*, if upon replacement of each of the variables by an arbitrary parameter λ times the variable, the function f is multiplied by λ^m. In symbols this means that

$$f(\lambda x_1, \lambda x_2, \cdots, \lambda x_n) = \lambda^m f(x_1, x_2, \cdots, x_n). \tag{26-1}$$

Thus,

$f(x, y, z) = x^2 - y^2 - 2z^2$ is homogeneous of degree 2,

$f(x, y) = \sqrt{y^2 - x^2} \sin^{-1} \dfrac{x}{y}$ is homogeneous of degree 1,

$f(x, y) = \dfrac{x^2}{y^2} \sin \dfrac{y}{x}$ is homogeneous of degree 0,

$f(x, y, z) = \sqrt{\dfrac{z}{x^2 + y^2}}$ is homogeneous of degree $-\dfrac{1}{2}$.

Theorem. *If $f(x_1, x_2, \cdots, x_n)$, with continuous partial derivatives, is homogeneous of degree m, then*

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_n \frac{\partial f}{\partial x_n} = mf(x_1, x_2, \cdots, x_n). \tag{26-2}$$

In order to prove this theorem, set

$$x_1' = \lambda x_1, \qquad x_2' = \lambda x_2, \cdots, x_n' = \lambda x_n.$$

Then, since f is homogeneous of degree m,

$$f(x_1', x_2', \cdots, x_n') = \lambda^m f(x_1, x_2, \cdots, x_n).$$

Differentiating with respect to λ one obtains

$$\frac{\partial f}{\partial x_1'} x_1 + \frac{\partial f}{\partial x_2'} x_2 + \cdots + \frac{\partial f}{\partial x_n'} x_n = m\lambda^{m-1} f(x_1, x_2, \cdots, x_n),$$

and setting $\lambda = 1$ gives the desired result (26-2) since

$$\frac{\partial f}{\partial x_k'} = \frac{\partial f}{\partial x_k}, \qquad (k = 1, 2, \cdots, n).$$

PROBLEM

Verify Euler's theorem for

(*a*) $f(x, y) = \sqrt{y^2 - x^2} \sin^{-1} \dfrac{y}{x}$;

(b) $f(x, y, z) = x^2y + xy^2 + 2xyz$;

(c) $f(x, y) = e^{x/y}$;

(d) $f(x, y) = \dfrac{\sqrt{x + y}}{y}$;

(e) $f(x, y) = \dfrac{x^2 + y^2}{x^2 - y^2}$.

27. Directional Derivatives. The relation expressed in (24-2) has an important special case when x and y are functions of the distance s along some curve C, which goes through the point (x, y). The curve C may be thought to be represented by a pair of parametric equations

$$x = x(s),$$
$$y = y(s),$$

where x and y are assumed to possess continuous derivatives with respect to the arc parameter s.

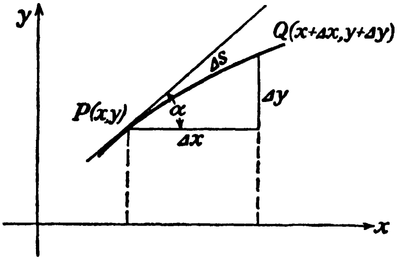

FIG. 16.

Let P (Fig. 16) be any point of the curve C at which $f(x, y)$ is defined and has partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$. Let

$$Q(x + \Delta x, y + \Delta y)$$

be a point close to P on this curve. If Δs is the length of the arc PQ and Δf is the change in f due to the increments Δx and Δy, then

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s}$$

gives the rate of change of f along C at the point (x, y). But

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds},$$

and

$$\frac{dx}{ds} = \lim_{\Delta s \to 0} \frac{\Delta x}{\Delta s} = \cos \alpha, \qquad \frac{dy}{ds} = \lim_{\Delta s \to 0} \frac{\Delta y}{\Delta s} = \sin \alpha.$$

Therefore,

$$\frac{df}{ds} = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \sin \alpha, \tag{27-1}$$

and it is evident that $\frac{df}{ds}$ depends on the direction of the curve. For this reason $\frac{df}{ds}$ is called the *directional derivative.* It represents the rate of change of f in the direction of the tangent to the particular curve chosen for the point (x, y). If $\alpha = 0$,

$$\frac{df}{ds} = \frac{\partial f}{\partial x},$$

which is the rate of change of f in the direction of the x-axis. If $\alpha = \frac{\pi}{2}$,

$$\frac{df}{ds} = \frac{\partial f}{\partial y},$$

which is the rate of change of f in the direction of the y-axis.

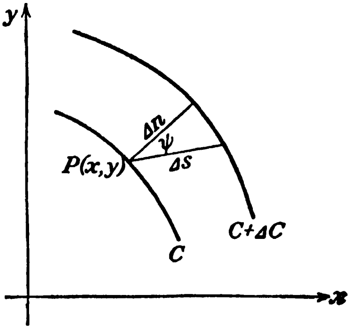

FIG. 17.

Let $z = f(x, y)$, which can be interpreted as the equation of a surface, be represented by drawing the contour lines on the xy-plane for various values of z. Let C (Fig. 17) be the curve in the xy-plane corresponding to the value $z = \gamma$, and let $C + \Delta C$ be the neighboring contour line for $z = \gamma + \Delta\gamma$. Then $\frac{\Delta f}{\Delta s} \equiv \frac{\Delta \gamma}{\Delta s}$ is the average rate of change of f with respect to the distance Δs between C and $C + \Delta C$. Apart from infinitesimals of higher order,

$$\frac{\Delta n}{\Delta s} = \cos \psi,$$

where Δn denotes the distance from C to $C + \Delta C$ along the normal to C at (x, y), and ψ is the angle between Δn and Δs; hence $\frac{dn}{ds} = \cos \psi$. Therefore

$$\frac{df}{ds} = \frac{df}{dn} \cdot \frac{dn}{ds} = \frac{df}{dn} \cos \psi. \tag{27-2}$$

This relation shows that the derivative of f in any direction may be found by multiplying the derivative along the normal by the cosine of the angle ψ between the particular direction and the normal. This derivative in the direction of the normal is called the *normal derivative* of f. Its numerical value obviously is the maximum value which $\frac{df}{ds}$ can take for any direction. In applied mathematics the vector in the direction of the normal, of magnitude $\frac{df}{dn}$, is called the *gradient.*

Example. Using (27-1) find the value of α which makes $\frac{df}{ds}$ a maximum, considering x and y to be fixed. Find the expression for this maximum value of $\frac{df}{ds}$.

Since $\frac{df}{ds} = f_x \cos \alpha + f_y \sin \alpha$,

$$\frac{d}{d\alpha}\left(\frac{df}{ds}\right) = -f_x \sin \alpha + f_y \cos \alpha.$$

The condition for a maximum requires that

$$\tan \alpha_1 = \frac{f_y}{f_x}, \qquad \text{or} \qquad \alpha_1 = \tan^{-1} \frac{f_y}{f_x}.$$

Using this value of α_1,

$$\frac{df}{dn} = f_x \frac{f_x}{\sqrt{f_x^2 + f_y^2}} + f_y \frac{f_y}{\sqrt{f_x^2 + f_y^2}} = \sqrt{f_x^2 + f_y^2}.$$

The relation (27-2) can be derived directly by use of this expression for $\frac{df}{dn}$. If α (Fig. 18) gives any direction different from the direction given by α_1, then

$$\frac{df}{ds} = f_x \cos \alpha + f_y \sin \alpha.$$

But $\alpha = \alpha_1 - \psi$, so that

$$\frac{df}{ds} = f_x(\cos\alpha_1 \cos\psi + \sin\alpha_1 \sin\psi) + f_y(\sin\alpha_1 \cos\psi - \cos\alpha_1 \sin\psi).$$

Since

$$\cos\alpha_1 = \frac{f_x}{\sqrt{f_x^2 + f_y^2}}$$

and

$$\sin\alpha_1 = \frac{f_y}{\sqrt{f_x^2 + f_y^2}},$$

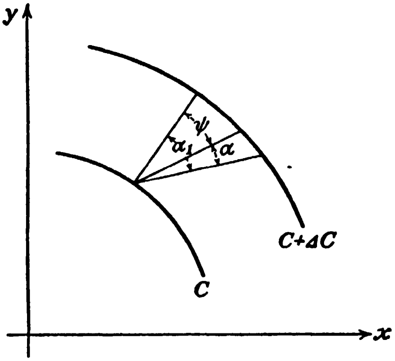

Fig. 18.

$$\begin{aligned}\frac{df}{ds} &= f_x\frac{f_x}{\sqrt{f_x^2 + f_y^2}}\cos\psi + f_x\frac{f_y}{\sqrt{f_x^2 + f_y^2}}\sin\psi \\ &\quad + f_y\frac{f_y}{\sqrt{f_x^2 + f_y^2}}\cos\psi - f_y\frac{f_x}{\sqrt{f_x^2 + f_y^2}}\sin\psi \\ &= \frac{f_x^2 + f_y^2}{\sqrt{f_x^2 + f_y^2}}\cos\psi = \sqrt{f_x^2 + f_y^2}\cos\psi \\ &= \frac{df}{dn}\cos\psi.\end{aligned}$$

PROBLEMS

1. Find the directional derivative of $f(x, y) = x^2y + \sin xy$ at $\left(1, \frac{\pi}{2}\right)$, in the direction of the line making an angle of 45° with the x-axis.

2. Find

$$\frac{df}{dn} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2},$$

if $x = r\cos\theta$, $y = r\sin\theta$, and f is a function of the variables r and θ.

3. Find the directional derivative of $f(x, y) = x^3y + e^{yx}$ in the direction of the curve which, at the point (1, 1), makes an angle of 30° with the x-axis.

4. Find the normal derivative of $f(x, y) = x^2 + y^2$.

28. Tangent Plane and Normal Line to a Surface. It will be recalled that

$$Ax + By + Cz = D$$

is the equation of a plane, where the coefficients A, B, and C are called the direction components of the normal to the plane.

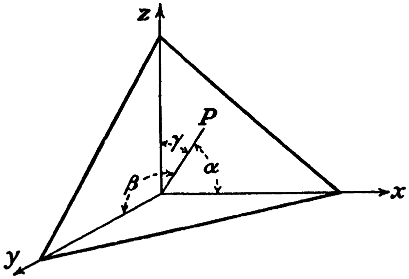

FIG. 19.

If α, β, and γ (Fig. 19) are the direction angles made by the normal to the plane from the origin, then

$$\cos \alpha = \frac{A}{\sqrt{A^2 + B^2 + C^2}}, \qquad \cos \beta = \frac{B}{\sqrt{A^2 + B^2 + C^2}},$$
$$\cos \gamma = \frac{C}{\sqrt{A^2 + B^2 + C^2}}.$$

Therefore,

$$\cos \alpha : \cos \beta : \cos \gamma = A : B : C.$$

If the plane passes through the point (x_0, y_0, z_0), its equation can be written as

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

There is also a normal form for the equation of a plane, entirely analogous to the normal form for the equation of the straight line in the plane. This form is

$$x \cos \alpha + y \cos \beta + z \cos \gamma = p,$$

or

$$\frac{A}{\sqrt{A^2+B^2+C^2}}x+\frac{B}{\sqrt{A^2+B^2+C^2}}y+\frac{C}{\sqrt{A^2+B^2+C^2}}z = \frac{D}{\sqrt{A^2+B^2+C^2}},$$

in which $p = \frac{D}{\sqrt{A^2+B^2+C^2}}$ is the distance from the origin to the plane.

Consider a surface defined by $z = f(x, y)$, in which x and y are considered as the independent variables. Then

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y. \tag{28-1}$$

If x_0 and y_0 are chosen, z_0 is determined by $z = f(x, y)$. Let $\Delta x = x - x_0$ and $\Delta y = y - y_0$, and denote dz by $z - z_0$. Then (28-1) becomes

$$z - z_0 = \left.\frac{\partial f}{\partial x}\right|_{(x_0,\, y_0)}(x - x_0) + \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}(y - y_0), \tag{28-2}$$

which is the equation of a plane. If this plane is cut by the plane $x = x_0$, the equation of the line of intersection is

$$z - z_0 = \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}(y - y_0),$$

and this is the tangent line to the curve $z = f(x_0, y)$ at the point (x_0, y_0, z_0). Similarly, the line of intersection of the plane defined by (28-2) and the plane $y = y_0$ is the tangent line to the curve $z = f(x, y_0)$ at (x_0, y_0, z_0). The plane defined by (28-2) is called the *tangent plane* to the surface

$$z = f(x, y)$$

at (x_0, y_0, z_0).

The direction cosines of the normal to this plane are proportional to

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,\, y_0)}, \quad \left.\frac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}, \quad -1.$$

The equation of the normal line to the plane (28-2) at (x_0, y_0, z_0) is therefore

$$(28\text{-}3)\qquad \frac{x - x_0}{\left.\dfrac{\partial f}{\partial x}\right|_{(x_0,\, y_0)}} = \frac{y - y_0}{\left.\dfrac{\partial f}{\partial y}\right|_{(x_0,\, y_0)}} = \frac{z - z_0}{-1}.$$

This line is defined as *the normal* to the surface at (x_0, y_0, z_0). Figure 20 shows the difference between $dz = RP'$ and $\Delta z = RQ$.

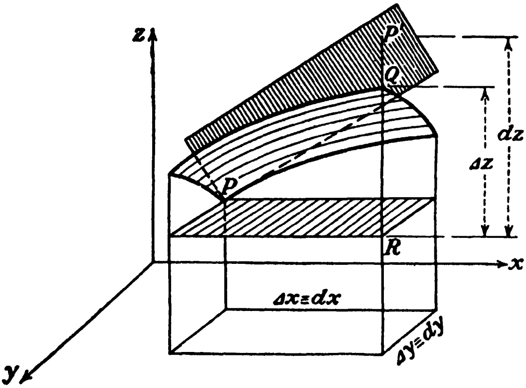

Fig. 20.

$P(x_0, y_0, z_0)$ is the point of tangency and $R(x_0 + \Delta x, y_0 + \Delta y, z_0)$ is in the plane $z = z_0$. PP' is the tangent plane.

In case the equation of the surface is given in the form

$$F(x, y, z) = 0,$$

the tangent plane and the normal line at (x_0, y_0, z_0) have the respective equations

$$(28\text{-}4)\qquad \left.\frac{\partial F}{\partial x}\right|_{(x_0,\, y_0,\, z_0)}(x - x_0) + \left.\frac{\partial F}{\partial y}\right|_{(x_0,\, y_0,\, z_0)}(y - y_0) + \left.\frac{\partial F}{\partial z}\right|_{(x_0,\, y_0,\, z_0)}(z - z_0) = 0$$

and

$$(28\text{-}5)\qquad \frac{x - x_0}{\left.\dfrac{\partial F}{\partial x}\right|_{(x_0,\, y_0,\, z_0)}} = \frac{y - y_0}{\left.\dfrac{\partial F}{\partial y}\right|_{(x_0,\, y_0,\, z_0)}} = \frac{z - z_0}{\left.\dfrac{\partial F}{\partial z}\right|_{(x_0,\, y_0,\, z_0)}}.$$

These equations follow directly from (25-4).

Example 1. At (6, 2, 3) on the surface $x^2 + y^2 + z^2 = 49$, the tangent plane has the equation

$$2x\Big|_{(6,2,3)}(x-6) + 2y\Big|_{(6,2,3)}(y-2) + 2z\Big|_{(6,2,3)}(z-3) = 0$$

or

$$6x + 2y + 3z = 49.$$

The normal line is

$$\frac{x-6}{12} = \frac{y-2}{4} = \frac{z-3}{6}.$$

Example 2. For (2, 1, 4) on the surface $z = x^2 + y^2 - 1$, the tangent plane is

$$z - 4 = 2x\Big|_{(2,1)}(x-2) + 2y\Big|_{(2,1)}(y-1)$$

or

$$4x + 2y - z = 6.$$

The normal line is

$$\frac{x-2}{4} = \frac{y-1}{2} = \frac{z-4}{-1}.$$

PROBLEMS

1. Find the distance from the origin to the plane $x + y + z = 1$.

2. Find the equations of the tangent plane and the normal line to

(*a*) $2x^2 + 3y^2 + 4z^2 = 6$ at $(1, 1, \frac{1}{2})$;

(*b*) $\dfrac{x^2}{4} + \dfrac{y^2}{9} - \dfrac{z^2}{16} = 1$ at (4, 3, 8);

(*c*) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ at (x_0, y_0, z_0).

3. Referring to (28-4), show that

$$\cos\alpha : \cos\beta : \cos\gamma = \frac{\partial F}{\partial x} : \frac{\partial F}{\partial y} : \frac{\partial F}{\partial z},$$

where $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are direction cosines of the normal line.

4. Show that the sum of the intercepts on the coordinate axes of any tangent plane to $x^{1/2} + y^{1/2} + z^{1/2} = a^{1/2}$ is constant.

29. Space Curves. It will be recalled that a plane curve C whose equation is

$$y = f(x) \tag{29-1}$$

can be represented in infinitely many ways by a pair of parametric equations

$$x = x(t),\quad y = y(t), \tag{29-2}$$

so chosen that when the independent variable t runs continuously through some set of values $t_1 \le t \le t_2$, the corresponding values of x and y determined by (29-2) satisfy (29-1).

For example, the equation of the upper half of a unit circle with the center at the origin of the cartesian system,

$$y = \sqrt{1 - x^2},$$

can be represented parametrically as

$$x = \cos t,\quad y = \sin t, \qquad (0 \le t \le \pi),$$

or

$$x = t,\quad y = \sqrt{1 - t^2}, \qquad (0 \le t \le 1),$$

or

$$x = 2t,\quad y = \sqrt{1 - 4t^2}, \qquad (0 \le t \le \tfrac{1}{2}).$$

Similarly, a space curve C can be represented by means of a set of equations

$$x = x(t),\quad y = y(t),\quad z = z(t), \tag{29-3}$$

so selected that when t runs through some set of values, the coordinates of the point $P(x, y, z)$, defined by (29-3), trace out the desired curve C.

It will be assumed that the functions in (29-2) and (29-3) possess continuous derivatives with respect to t, which implies that the curve C has a continuously turning tangent as the point P moves along the curve.

Let $P(x_0, y_0, z_0)$ (Fig. 21) be a point of the curve C defined by (29-3) that corresponds to some value t_0 of the parameter t, and let Q be the point $(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$ that corresponds to $t = t_0 + \Delta t$. The direction ratios of the line PQ joining P and Q are

$$\frac{\Delta x}{\Delta c}:\frac{\Delta y}{\Delta c}:\frac{\Delta z}{\Delta c} = \frac{\Delta x}{\Delta t}:\frac{\Delta y}{\Delta t}:\frac{\Delta z}{\Delta t}.$$

If Δt is allowed to approach zero, Δx, Δy, and Δz all tend to zero, so that the direction ratios of the tangent line at $P(x_0, y_0, z_0)$ are proportional to $\left(\frac{dx}{dt}\right)_{t=t_0}:\left(\frac{dy}{dt}\right)_{t=t_0}:\left(\frac{dz}{dt}\right)_{t=t_0}$. Hence, the equation of the tangent line to C at P is

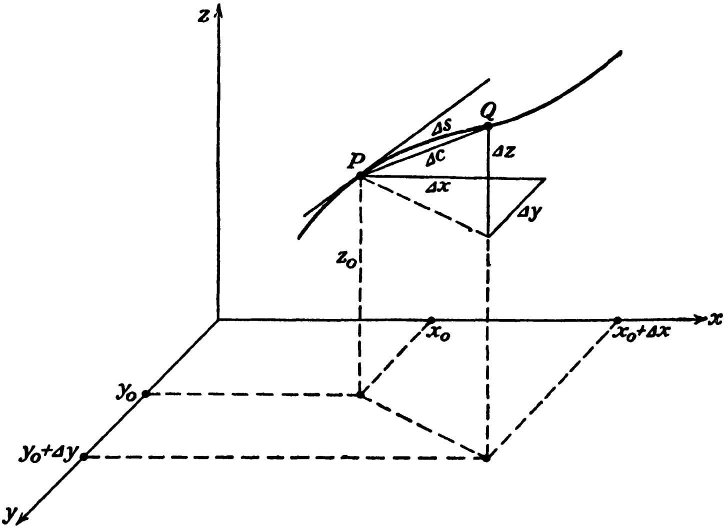

Fig. 21.

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} = \frac{z - z_0}{z'(t_0)},$$

where primes denote derivatives with respect to t.

Example. The equation of the tangent line to the circular helix

$$x = a \cos t,$$
$$y = a \sin t,$$
$$z = at.$$

at $t = \frac{\pi}{6}$, is

$$\frac{x - \frac{\sqrt{3}}{2}a}{-\frac{a}{2}} = \frac{y - \frac{a}{2}}{\frac{\sqrt{3}}{2}a} = \frac{z - \frac{\pi a}{6}}{a}.$$

The element of arc ds is given by

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2,$$

so that the length of a space curve C can be calculated from

$$L = \int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt.$$

The length of the part of the helix between the points $(a, 0, 0)$ and $\left(0, a, \frac{\pi a}{2}\right)$ is

$$L = \frac{\sqrt{2}}{2}\pi a.$$

30. Directional Derivatives in Space. There is no essential difficulty in extending the results of Sec. 27 to any number of variables. Thus, if $u = f(x, y, z)$ is a suitably restricted function of the independent variables x, y, and z, then the directional derivative along a space curve whose tangent line at some point $P(x, y, z)$ (Fig. 21) has the direction cosines $\cos(x, s)$, $\cos(y, s)$, and $\cos(z, s)$ is

$$\begin{aligned} \frac{du}{ds} &= \frac{\partial u}{\partial x}\cos(x, s) + \frac{\partial u}{\partial y}\cos(y, s) + \frac{\partial u}{\partial z}\cos(z, s) \\ &= \frac{\partial u}{\partial x}\frac{dx}{ds} + \frac{\partial u}{\partial y}\frac{dy}{ds} + \frac{\partial u}{\partial z}\frac{dz}{ds}. \end{aligned}$$

The magnitude of the normal derivative is given by

$$\frac{du}{dn} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}. \tag{30-1}$$

PROBLEMS

1. Find the equation of the tangent line to the helix

$$x = a\cos t, \qquad y = a\sin t, \qquad z = at,$$

at the point where $t = \frac{\pi}{4}$. Find the length of the helix between the points $t = 0$ and $t = \frac{\pi}{4}$.

2. Find the directional derivative of $f = xyz$ at $(1, 2, 3)$ in the direction of the line that makes equal angles with the coordinate axes.

3. Find the normal derivative of $f = x^2 + y^2 + z^2$ at $(1, 2, 3)$.

4. Show that the square root of the sum of the squares of the directional derivatives in three perpendicular directions is equal to the normal derivative.

5. Express the normal derivative (30-1) in spherical and cylindrical coordinates, for which the equations of transformation are

(a) $x = r \sin\theta \cos\varphi,\ y = r \sin\theta \sin\varphi,\ z = r\cos\theta$;
(b) $x = r\sin\theta,\ y = r\cos\theta,\ z = z$.

31. Partial Derivatives of Higher Order. The first partial derivatives of a function $f(x_1, x_2, \cdots, x_n)$ are functions of $x_1, x_2, \cdots, x_n$ and they may have derivatives with respect to some or all of these variables. These derivatives are called *second partial derivatives.*

If there are only two variables x and y the function $u = f(x, y)$ may have the second partial derivatives

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial x} \equiv \frac{\partial^2 f}{\partial x^2} \equiv f_{xx},$$
$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} \equiv \frac{\partial^2 f}{\partial y\,\partial x} \equiv f_{xy},$$
$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y} \equiv \frac{\partial^2 f}{\partial x\,\partial y} \equiv f_{yx},$$
$$\frac{\partial}{\partial y}\frac{\partial f}{\partial y} \equiv \frac{\partial^2 f}{\partial y^2} \equiv f_{yy}.$$

It should be noted carefully that f_{xy} $\left(\text{or } \frac{\partial^2 f}{\partial y\,\partial x}\right)$ means that $\frac{\partial f}{\partial x}$ is found first and the result is differentiated with respect to y. Thus $\frac{\partial^2 f}{\partial y\,\partial x}$ and $\frac{\partial^2 f}{\partial x\,\partial y}$ differ in the order in which the differentiation is performed, and, consequently, the results may be different. The following theorem, due to Schwarz,* states sufficient conditions for inversion of the order of differentiation.

Theorem. *If f_x, f_y, and f_{xy} exist in the vicinity of the point (x, y) and if f_{xy} is continuous at (x, y), then f_{yx} exists at this point and is identical with f_{xy}.*

Form the function

$$\varphi(x) \equiv f(x, y + k) - f(x, y), \tag{31-1}$$

---

* De la Vallée Poussin, C. J., Cours d'analyse infinitésimale. vol. 1, p. 146.

which for a fixed y and k satisfies the conditions of the mean-value theorem. Hence, making use of (20-4),

$$\begin{aligned}\varphi(x+h)-\varphi(x) &= h\varphi_x(x+\theta_1 h)\\ &= h[f_x(x+\theta_1 h,\, y+k)-f_x(x+\theta_1 h,\, y)],\end{aligned}$$

where $0 < \theta_1 < 1$.
The expression in the brackets is a function of y to which the mean-value theorem is applicable, so that

$$f_x(x+\theta_1 h,\, y+k)-f_x(x+\theta_1 h,\, y) = kf_{xy}(x+\theta_1 h,\, y+\theta_2 k).$$

Since f_{xy} is assumed to be continuous at (x, y),

$$f_{xy}(x+\theta_1 h,\, y+\theta_2 k) = f_{xy}(x, y)+\epsilon,$$

where

$$\lim_{\substack{h\to 0\\ k\to 0}} \epsilon = 0.$$

Therefore,

$$\varphi(x+h)-\varphi(x) = hk[f_{xy}(x, y)+\epsilon]. \tag{31-2}$$

From the defining equation (31-1) it is seen that (31-2) is equivalent to

$$\begin{aligned}f(x+h,\, y+k)-f(x+h,\, y)-f(x,\, y+k)+f(x, y)&\\ = hk[f_{xy}(x, y)+\epsilon]&.\end{aligned}$$

Dividing through by k gives

$$\begin{aligned}\frac{f(x+h,\, y+k)-f(x+h,\, y)}{k}-\frac{f(x,\, y+k)-f(x, y)}{k}&\\ = h[f_{xy}(x, y)+\epsilon]&.\end{aligned}$$

Passing to the limit as $k \to 0$, one obtains

$$f_y(x+h,\, y)-f_y(x, y) = hf_{xy}(x, y)+h\lim_{k\to 0}\epsilon,$$

or

$$\frac{f_y(x+h,\, y)-f_y(x, y)}{h} = f_{xy}(x, y)+\lim_{k\to 0}\epsilon.$$

The limit of this expression as $h \to 0$ is

$$f_{yx}(x, y) = f_{xy}(x, y).$$

Thus the order of differentiation is immaterial if the assumptions regarding the partial derivatives are satisfied. These conditions certainly will be satisfied if it is known that f_x and f_y

possess derivatives f_{xy} and f_{yx} which are continuous functions. The theorem is true under less restrictive conditions, but the detailed statement of these conditions is so involved that it will not be given here.*

From the result just established it is clear that

$$f_{xxy}(x, y) = f_{yxx}(x, y),$$

if these partial derivatives are continuous. For

$$f_{xxy}(x, y) \equiv \frac{\partial^2 f_x(x, y)}{\partial y\, \partial x} = \frac{\partial^2 f_x(x, y)}{\partial x\, \partial y} \equiv f_{xyx}(x, y) = f_{yxx}(x, y).$$

32. Higher Derivatives of Implicit Functions. The problem of calculating the derivative of y with respect to x when y is an implicit function of the independent variable x defined by

$$f(x, y) = 0, \tag{32-1}$$

was discussed in Sec. 24. It was shown there that

$$f_x(x, y) + f_y(x, y) \frac{dy}{dx} = 0. \tag{32-2}$$

Differentiating this equation again and assuming that all the derivatives involved are continuous functions of x and y, gives

$$f_{xx}(x, y) + 2f_{xy}(x, y) \frac{dy}{dx} + f_{yy}(x, y)\left(\frac{dy}{dx}\right)^2 + f_y(x, y) \frac{d^2y}{dx^2} = 0. \tag{32-3}$$

If $f_y(x, y) \neq 0$ at the point where the derivative is desired, (32-3) can be solved for $\frac{d^2y}{dx^2}$ and the value of $\frac{dy}{dx}$ substituted from (32-2). The result is

$$\frac{d^2y}{dx^2} = -\frac{f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2}{f_y^3}.$$

The process can be continued to obtain the derivatives of higher orders.

A similar procedure can be employed to calculate the partial derivatives of a function z of two independent variables x and y defined implicitly by an equation of the form

$$f(x, y, z) = 0. \tag{32-4}$$

---

* See E. W. HOBSON, Theory of Functions of a Real Variable, vol. 1, pp. 424–429.

Differentiating (32-4) with respect to x and y in turn gives

$$\begin{aligned} &f_x(x, y, z) + f_z(x, y, z)\frac{\partial z}{\partial x} = 0, \\ (32\text{-}5) \qquad &f_y(x, y, z) + f_z(x, y, z)\frac{\partial z}{\partial y} = 0. \end{aligned}$$

If $f_z(x, y, z)$ does not vanish for those values of x, y, and z that satisfy (32-4), then Eqs. (32-5) can be solved for $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$. Partial derivatives of higher order can then be obtained by differentiating equations (32-5).

Example. Let it be required to find the derivatives of second order of the function z defined implicitly by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Differentiating this equation with respect to x and y gives

$$\begin{aligned} &\frac{2x}{a^2} + \frac{2z}{c^2}\frac{\partial z}{\partial x} = 0, \\ (32\text{-}6) \qquad &\frac{2y}{b^2} + \frac{2z}{c^2}\frac{\partial z}{\partial y} = 0. \end{aligned}$$

Differentiating the first of Eqs. (32-6) with respect to x and y, one obtains

$$\frac{2}{a^2} + \frac{2}{c^2}\left(\frac{\partial z}{\partial x}\right)^2 + \frac{2z}{c^2}\frac{\partial^2 z}{\partial x^2} = 0,$$

$$\frac{2}{c^2}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + \frac{2z}{c^2}\frac{\partial^2 z}{\partial x\,\partial y} = 0.$$

Solving for $\frac{\partial^2 z}{\partial x^2}$ and $\frac{\partial^2 z}{\partial x\,\partial y}$ and making use of (32-6), one obtains

$$\frac{\partial^2 z}{\partial x^2} = -\frac{c^2}{a^4}\frac{a^2z^2 + c^2x^2}{z^3},$$

$$\frac{\partial^2 z}{\partial x\,\partial y} = -\frac{c^4}{a^2b^2}\frac{xy}{z^3}.$$

In a similar way the differentiation of the second of Eqs. (32-6) with respect to y yields

$$\frac{\partial^2 z}{\partial y^2} = -\frac{c^2}{b^4}\frac{b^2z^2 + c^2y^2}{z^3}.$$

PROBLEMS

1. Verify that $\dfrac{\partial^2 f}{\partial x\,\partial y} = \dfrac{\partial^2 f}{\partial y\,\partial x}$ for

   (a) $f = \cos xy^2$;
   (b) $f = \sin^2 x \cos y$;
   (c) $f = e^{y/x}$;
   (d) $f = (x^2 + y^2)^{\frac{1}{2}}$;
   (e) $f = (x^2 - 2y)^2 + \sqrt{xy}$.

2. Prove that if

   (a) $f(x, y) = \log (x^2 + y^2) + \tan^{-1} \dfrac{y}{x}$, then $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0$;

   (b) $f(x, y, z) = (x^2 + y^2 + z^2)^{-\frac{1}{2}}$, then $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial^2 f}{\partial z^2} = 0$.

3. Find y', y'', y''', if $x^3 + y^3 - 3axy = 0$.

4. Find $\dfrac{\partial z}{\partial x}$, $\dfrac{\partial z}{\partial y}$, $\dfrac{\partial^2 z}{\partial x^2}$, $\dfrac{\partial^2 z}{\partial x\,\partial y}$, and $\dfrac{\partial^2 z}{\partial y^2}$ at $(1, 1, 1)$, if $x^2 - y^2 + z^2 = 1$.

5. Find $\dfrac{\partial z}{\partial x}$, if

   (a) $xz^2 - yz^2 + xy^2z - 5 = 0$;
   (b) $xz^3 - yz + 3xy = 0$.

33. Change of Variables. In a great variety of problems in analysis it is required to express the derivatives of a given function of one set of variables in terms of another set of variables. Some of the simpler problems of this sort were discussed in Secs. 24 and 25. This section contains a formal treatment of several of the more difficult of such problems. The question of the existence of solutions of the equations to be considered in this section will be left open until Chap. XII, where the problem of the differentiation of implicit functions will be reexamined in the light of the existence theorems established there.

The sole purpose of the present section is to develop manipulative skill in calculating the derivatives of implicit functions and to indicate the formal modes of attack on the problem. The continuity of the functions and their partial derivatives is assumed throughout this section and will not be referred to again.

Let

$$w = f(u, v) \tag{33-1}$$

denote a function of two independent variables u and v, and suppose that u and v are connected with some other variables x and y by means of the relations

$$\text{(33-2)} \qquad \begin{cases} x = x(u, v), \\ y = y(u, v). \end{cases}$$

If the equations (33-2) are solved for x and y to yield

$$\text{(33-3)} \qquad \begin{cases} u = u(x, y), \\ v = v(x, y), \end{cases}$$

and the expressions (33-3) are substituted in (33-1) for u and v, there will result a function of x and y, say,

$$\text{(33-4)} \qquad w = F(x, y).$$

The partial derivatives of w with respect to x and y can be calculated from (33-4) directly, but frequently it is impracticable to effect the solution (33-3), and it is desirable to consider an indirect mode of calculation. By the rule for the differentiation of composite functions,

$$\text{(33-5)} \qquad \begin{cases} \dfrac{\partial w}{\partial u} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial u}, \\ \dfrac{\partial w}{\partial v} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial v}. \end{cases}$$

The partial derivatives $\dfrac{\partial x}{\partial u}$, $\dfrac{\partial y}{\partial u}$, $\dfrac{\partial x}{\partial v}$, and $\dfrac{\partial y}{\partial v}$ can be calculated from (33-2), and hence they may be regarded as known functions of u and v. The partial derivatives in the left-hand members of (33-5) are also known functions of u and v since they can be calculated from (33-1).

Hence, equations (33-5) may be regarded as linear equations for the determination of $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$. Assuming that

$$J(u, v) \equiv \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} \neq 0,$$

and solving by Cramer's rule gives

$$\frac{\partial w}{\partial x} = \frac{\begin{vmatrix} \dfrac{\partial w}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial w}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}}{J(u, v)}, \qquad \frac{\partial w}{\partial y} = \frac{\begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial w}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial w}{\partial v} \end{vmatrix}}{J(u, v)}.$$

The resulting expressions for $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ are known functions of u and v and thus can be treated exactly like (33-1) if it is desirable to calculate the derivatives of higher orders.

As an example, consider the function $w(r, \theta)$, and let it be required to calculate the partial derivatives of w with respect to x and y, where $x = r\cos\theta$ and $y = r\sin\theta$. Now

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial w}{\partial x}\cos\theta + \frac{\partial w}{\partial y}\sin\theta,$$

$$\frac{\partial w}{\partial \theta} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial \theta} = -\frac{\partial w}{\partial x}r\sin\theta + \frac{\partial w}{\partial y}r\cos\theta.$$

Solving these equations for $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ in terms of $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial \theta}$ gives

$$\frac{\partial w}{\partial x} = \cos\theta\,\frac{\partial w}{\partial r} - \frac{\sin\theta}{r}\frac{\partial w}{\partial \theta},$$

$$\frac{\partial w}{\partial y} = \sin\theta\,\frac{\partial w}{\partial r} + \frac{\cos\theta}{r}\frac{\partial w}{\partial \theta}.$$

The function J is, in this case,

$$\begin{vmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{vmatrix} = r,$$

which does not vanish unless $r = 0$.

As a somewhat more complicated instance of implicit differentiation, consider a pair of equations

$$\text{(33-6)} \qquad \begin{cases} F(x, y, u, v) = 0, \\ G(x, y, u, v) = 0, \end{cases}$$

and let it be supposed that they can be solved for u and v in terms of x and y to yield

$$\text{(33-7)} \qquad \begin{cases} u = u(x, y), \\ v = v(x, y). \end{cases}$$

The partial derivatives of u and v with respect to x and y can be obtained in the following manner. Considering x and y as the independent variables and differentiating equations (33-6) with respect to x and y gives

$$(33\text{-}8)\quad \begin{cases} \dfrac{\partial F}{\partial x} + \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial x} = 0; & \dfrac{\partial F}{\partial y} + \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial y} = 0; \\ \dfrac{\partial G}{\partial x} + \dfrac{\partial G}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial G}{\partial v}\dfrac{\partial v}{\partial x} = 0; & \dfrac{\partial G}{\partial y} + \dfrac{\partial G}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial G}{\partial v}\dfrac{\partial v}{\partial y} = 0. \end{cases}$$

Equations (33-8) are linear in $\dfrac{\partial u}{\partial x}$, $\dfrac{\partial u}{\partial y}$, $\dfrac{\partial v}{\partial x}$, and $\dfrac{\partial v}{\partial y}$.

As a consequence of the hypothesis that Eqs. (33-6) possess the solution (33-7), it follows that

$$J(u, v) \equiv \begin{vmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{vmatrix} \neq 0$$

in the region for which the solution (33-7) is valid.* Accordingly, the partial derivatives in question can be determined from (33-8) by Cramer's rule.

A special case of Eqs. (33-6) is interesting. Let

$$\begin{aligned} x &= f(u, v), \\ y &= g(u, v). \end{aligned}$$

Differentiating these equations with respect to x and remembering that x and y are independent variables, one obtains

$$(33\text{-}9)\quad \begin{cases} 1 = \dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial x}, \\ 0 = \dfrac{\partial g}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial g}{\partial v}\dfrac{\partial v}{\partial x}. \end{cases}$$

These equations can be solved for $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial v}{\partial x}$, if

$$J(u, v) \equiv \begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\ \dfrac{\partial g}{\partial u} & \dfrac{\partial g}{\partial v} \end{vmatrix} \neq 0.$$

Example 1. Let

$$\begin{aligned} u^2 - v^2 + 2x &= 0, \\ uv - y &= 0. \end{aligned}$$

---

* The proof of this assertion is given in Chap. XII.

Differentiating with respect tc x,

$$2\left(u\frac{\partial u}{\partial x} - v\frac{\partial v}{\partial x} + 1\right) = 0,$$
$$v\frac{\partial u}{\partial x} + u\frac{\partial v}{\partial x} = 0.$$

Hence

$$\frac{\partial u}{\partial x} = -\frac{u}{u^2+v^2}, \qquad \frac{\partial v}{\partial x} = \frac{v}{u^2+v^2}.$$

Differentiating the first of these results with respect to x gives

$$\frac{\partial^2 u}{\partial x^2} = \frac{-\frac{\partial u}{\partial x}(u^2+v^2) + 2\left(u\frac{\partial u}{\partial x} + v\frac{\partial v}{\partial x}\right)u}{(u^2+v^2)^2}$$
$$= \frac{u(u^2+v^2) - 2u(u^2-v^2)}{(u^2+v^2)^3} = \frac{u(3v^2-u^2)}{(u^2+v^2)^3}.$$

One obtains similarly $\frac{\partial^2 v}{\partial x^2}$, $\frac{\partial^2 u}{\partial x\,\partial y}$, and higher derivatives.

Example 2. Let

$$(a) \qquad \begin{cases} x = u + v, \\ y = 3u + 2v. \end{cases}$$

Differentiating with respect to x,

$$1 = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x},$$
$$0 = 3\frac{\partial u}{\partial x} + 2\frac{\partial v}{\partial x},$$

so that

$$\frac{\partial u}{\partial x} = -2, \qquad \frac{\partial v}{\partial x} = 3.$$

It is easily checked that

$$\frac{\partial u}{\partial y} = 1, \qquad \frac{\partial v}{\partial y} = -1.$$

Equations (a) can be solved for u and v in terms of x and y, and the result is

$$u = -2x + y,$$
$$v = 3x - y.$$

Regarding u and v as the independent variables and differentiating these equations with respect to u, one finds

$$1 = -2\frac{\partial x}{\partial u} + \frac{\partial y}{\partial u},$$

$$0 = 3\frac{\partial x}{\partial u} - \frac{\partial y}{\partial u}.$$

Hence,

$$\frac{\partial x}{\partial u} = 1, \qquad \frac{\partial y}{\partial u} = 3.$$

It should be noted that $\frac{\partial u}{\partial x}$ and $\frac{\partial x}{\partial u}$ are not reciprocals in general, as, of course, is obvious from (33-9).

Example 3. If $w = uv$ and

$$(b) \qquad \begin{cases} u^2 + v + x = 0, \\ v^2 - u - y = 0, \end{cases}$$

one can obtain $\frac{\partial w}{\partial x}$ as follows: Differentiation of w with respect to x gives

$$\frac{\partial w}{\partial x} = u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial x}.$$

The values of $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$ can be calculated from (b) as was done in Example 1. The reader will check that

$$\frac{\partial w}{\partial x} = -\frac{u + 2v^2}{1 + 4uv}, \qquad \frac{\partial w}{\partial y} = \frac{2u^2 - v}{1 + 4uv}.$$

PROBLEMS

1. If $u^2 + v^2 + y^2 - 2x = 0$, $u^3 + v^3 - x^3 + 3y = 0$, find $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial x}$, $\frac{\partial u}{\partial y}$, and $\frac{\partial v}{\partial y}$.

2. Find $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$, if $w = \frac{u}{v}$,

and
$$\begin{cases} x = u + v, \\ y = 3u + 2v. \end{cases}$$

3. Show that if $f(x, y, z) = 0$, then $\dfrac{\partial z}{\partial x}\dfrac{\partial x}{\partial z} = 1$ and $\dfrac{\partial x}{\partial y}\dfrac{\partial y}{\partial z}\dfrac{\partial z}{\partial x} = -1$. Note that in general $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial x}{\partial z}$ are not reciprocals.

4. If $x = x(u, v)$, $y = y(u, v)$, and $\dfrac{\partial x}{\partial u} = \dfrac{\partial y}{\partial v}$, and $\dfrac{\partial x}{\partial v} = -\dfrac{\partial y}{\partial u}$, then

$$\frac{\partial^2 w}{\partial u^2} + \frac{\partial^2 w}{\partial v^2} = \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right)\left[\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial x}{\partial v}\right)^2\right].$$

5. Show that the expressions

$$V_1 = \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2,$$

and

$$V_2 = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2},$$

upon change of variable by means of

$$x = r\cos\theta,$$
$$y = r\sin\theta,$$

become

$$V_1 = \left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2,$$

and

$$V_2 = \frac{\partial^2 z}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} + \frac{1}{r}\frac{\partial z}{\partial r}.$$

6. Show that

$$\frac{\partial^2 V}{\partial t^2} = c^2\frac{\partial^2 V}{\partial x^2},$$

if $V = f(x + ct) + g(x - ct)$, where f and g are any functions possessing continuous second derivatives.

7. Show that

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = e^{-2r}\left(\frac{\partial^2 V}{\partial r^2} + \frac{\partial^2 V}{\partial \theta^2}\right),$$

if $x = e^r\cos\theta$, $y = e^r\sin\theta$.

8. Find $\dfrac{\partial u}{\partial x}$, if

$$u^2 - v^2 - x^3 + 3y = 0,$$
$$u + v - y^2 - 2x = 0.$$

9. Prove that

$$\frac{\partial u}{\partial x}\frac{\partial y}{\partial u}+\frac{\partial v}{\partial x}\frac{\partial y}{\partial v}=0,$$

if $F(x, y, u, v) = 0$ and $G(x, y, u, v) = 0$.

10. If $V_1(x, y, z)$ and $V_2(x, y, z)$ satisfy the equation

$$\nabla^2 V \equiv \frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}+\frac{\partial^2 V}{\partial z^2}=0,$$

then

$$U \equiv V_1(x, y, z) + (x^2 + y^2 + z^2)V_2(x, y, z)$$

satisfies the equation

$$\nabla^2 \nabla^2 U = 0,$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}.$$

CHAPTER IV

DEFINITE INTEGRALS

34. Riemann Integral. The concept of a definite integral had its origin in problems dealing with the determination of the area under a curve. Despite the fact that the refined concept of what is known as the Riemann integral is very far removed from the geometrical notions that led to its definition, it seems desirable to begin the study of Riemann integration by presenting a reasonably careful definition of the definite integral based on the intuitive concept of the area under the curve. An analytical definition of the Riemann integral is given in the next section.

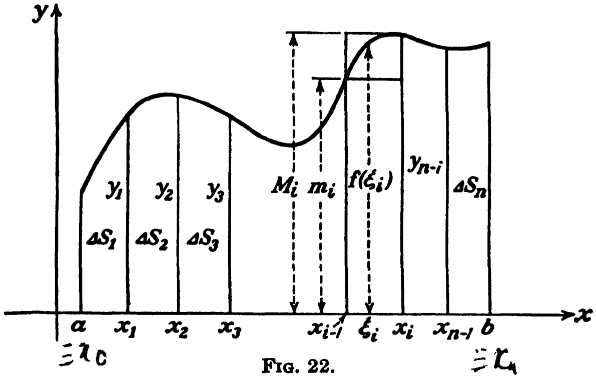

FIG. 22.

The function $f(x)$ considered in this section is assumed to be continuous and single-valued. Let the interval of definition of the function be $a \leq x \leq b$, and assume that the graph of the function lies above the x-axis. The area bounded by the curve $y = f(x)$, the x-axis, and the lines $x = a$ and $x = b$ will be denoted by S (Fig. 22). Divide the interval (a, b) into n parts by the points

$$a \equiv x_0, x_1, x_2, \cdots, x_{n-1}, x_n \equiv b,$$

whose abscissas satisfy the inequality

$$x_i < x_{i+1}, \qquad (i = 0, 1, 2, \cdots, n - 1),$$

and draw through these points of subdivision a set of ordinates

$$y_1, y_2, \cdots, y_{n-1}.$$

The area S will be divided into n vertical strips, the ith one of which has a base of length $\Delta x_i \equiv x_i - x_{i-1}$ and an area ΔS_i.

Denote the maximum and the minimum values of the function $y = f(x)$ in the interval Δx_i by M_i and m_i, respectively, and consider the rectangular areas

$$m_i \, \Delta x_i \qquad \text{and} \qquad M_i \, \Delta x_i.$$

Clearly,

$$m_i \, \Delta x_i \leq \Delta S_i \leq M_i \, \Delta x_i,$$

so that the area S is not less than

$$s_n \equiv \sum_{i=1}^{n} m_i \, \Delta x_i, \tag{34-1}$$

which represents the sum of the areas of the inscribed rectangles. On the other hand, the area S is not greater than

$$S_n \equiv \sum_{i=1}^{n} M_i \, \Delta x_i, \tag{34-2}$$

which stands for the sum of the areas of the circumscribing rectangles. Therefore, one can write

$$s_n \leq S \leq S_n. \tag{34-3}$$

The difference between (34-2) and (34-1) is

$$S_n - s_n = \sum_{i=1}^{n} (M_i - m_i) \, \Delta x_i, \tag{34-4}$$

and if the number n of subdivisions is increased indefinitely in such a way that all of the subintervals $\Delta x_i \to 0$, the difference $M_i - m_i$ will also tend to zero, since $f(x)$ is continuous. It does not follow without proof that $S_n - s_n$ will tend to zero, since the number of terms in the sum increases with the increase in n. It will be shown, however, that this is the case.

Consider the set of numbers

$$M_i - m_i, \qquad (i = 1, 2, \cdots, n),$$

associated with some particular mode of subdivision of (a, b) into

n parts, and denote the ~~greatest~~ of these numbers by d_n. Then the sum in (34-4) certainly will not be decreased if each $M_i - m_i$ is replaced by d_n, so that

$$S_n - s_n \leq d_n \sum_{i=1}^{n} \Delta x_i = d_n(b - a).$$

But $S_n - s_n$ is nonnegative, so that

$$0 \leq S_n - s_n \leq d_n(b - a),$$

and since $d_n \to 0$ when $n \to \infty$, it follows that

$$\lim_{n \to \infty} (S_n - s_n) = 0,$$

or

$$\lim_{n \to \infty} S_n = \lim_{n \to \infty} s_n.$$

The area S, under the curve $y = f(x)$, as is seen from (34-3), is intermediate to s_n and S_n, and since both s_n and S_n have the same limit, it is clear that their common limit is S. Thus,

$$\boxed{\lim_{n \to \infty} s_n = \lim_{n \to \infty} S_n = S.} \checkmark$$

Moreover, let ξ_i be any point in the interval Δx_i, and form the sum

$$S'_n = \sum_{i=1}^{n} f(\xi_i)\, \Delta x_i,$$

which represents the sum of the areas of rectangular strips whose heights are intermediate to m_i and M_i, so that

$$s_n \leq S'_n \leq S_n.$$

Since

$$\lim_{n \to \infty} s_n = \lim_{n \to \infty} S_n = S,$$

it follows that

$$\lim_{n \to \infty} S'_n = S.$$

Hence, the sum S'_n may be constructed for an arbitrary choice of the point ξ_i in the subinterval Δx_i, and its limit will be the same. This leads to the following definition:

Definition. *Let $f(x)$ be a continuous function defined in the interval $a \leq x \leq b$. Let the interval (a, b) be divided into n sub-*

intervals by inserting the points of subdivision x_i in such a way that

$$a \equiv x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n \equiv b.$$

Let ξ_i be any point in the interval of length $\Delta x_i = x_i - x_{i-1}$. The limit of the sum

$$\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i,$$

as $n \to \infty$ in such a way that the greatest $\Delta x_i \to 0$, is called the Riemann definite integral of $f(x)$ between the limits a and b. It is denoted by the symbol

$$\int_a^b f(x)\,dx.$$

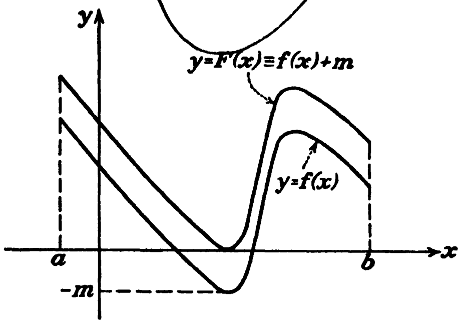

FIG. 23.

The geometrical reasoning that led to the definition of the integral of Riemann involved an assumption that $f(x)$ is a nonnegative function. The fact that this assumption is not an essential one can be seen from the following considerations. Assume first that $f(x)$ is negative throughout the interval (a, b). Then each term in the sum

$$\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$$

is negative, and the limit of the sum will be a negative number. Hence, areas lying below the x-axis must be reckoned as negative.

If $f(x)$ is not always of the same sign, denote the smallest value of $f(x)$ in (a, b) by $-m$ (Fig. 23). Then the function

$$F(x) = f(x) + m$$

is positive or zero throughout the interval (a, b), and one can form the integral of $F(x)$, namely,

$$\begin{aligned}\int_a^b F(x)\,dx &= \int_a^b [f(x) + m]\,dx \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} [f(\xi_i) + m]\,\Delta x_i \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i + \lim_{n\to\infty} \sum_{i=1}^{n} m\,\Delta x_i.\end{aligned}$$

The second term in the right-hand member of this equation gives $m(b - a)$, and if the integral of $F(x)$ exists,

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$$

will surely exist.

The restriction that a be less than b likewise can be removed. Thus, suppose that $a > b$, and let the interval (b, a) be subdivided into n parts by the points x_i such that

$$a \equiv x_0 > x_1 > x_2 > \cdot\cdot\cdot > x_{n-1} > x_n \equiv b.$$

Then the difference

$$x_i - x_{i-1} = \Delta x_i$$

is negative, so that the value of the sum

$$\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$$

will differ only in sign from the corresponding sum appearing in the definition. Hence,

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

In view of this relation (as well as from the geometrical interpretation of the definite integral), it appears sensible to say that

$$\int_a^a f(x)\,dx = 0.$$

It should be noted that the value of the definite integral $\int_a^b f(x)\,dx$ is a number S which certainly does not depend on the choice of the letter used to denote the variable of integration. Consequently, the letter denoting the variable of integration can be changed at will. Thus,

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt.$$

The foregoing discussion was restricted to the consideration of continuous functions only. However, the limit of the sum, figuring in the definition, may exist and be independent of the mode of subdivision of the interval (a, b) into subintervals Δx_i (as well as of the choice of ξ_i), even if $f(x)$ is not continuous.* Any function for which the integral of Riemann exists is said to be *integrable in the sense of Riemann.*

It is left to the reader to show:

(*a*) *That whatever be the choice of the three numbers a, b, and c,*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx;$$

(*b*) *That if $f_1(x)$ and $f_2(x)$ are two integrable functions, then*

$$\int_a^b [f_1(x) + f_2(x)]\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx;$$

(*c*) *That if $f(x)$ is an integrable function and c is any constant, then*

$$\int_a^b c\,f(x)\,dx = c\int_a^b f(x)\,dx;$$

(*d*) *That if $|f(x)|$ denotes the absolute value of the integrable function $f(x)$, then*

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b \left|f(x)\right|\,dx, \qquad \text{if} \qquad a < b.$$

35. Riemann Integral (*Continued*). The existence of the common limit S of the sums s_n and S_n in Sec. 34 hinged upon the geometrical concept of the area bounded by the curve $y = f(x)$, the ordinates $x = a$ and $x = b$, and the x-axis. A geometrical argument does seem convincing enough until one begins to inquire carefully into the possibility of graphical representation of continuous functions. It was indicated above that one need not go very far out of his way to meet some continuous functions which cannot be represented graphically. The function

---

* See Sec. 35.

$$\begin{aligned} f(x) &= x \sin \frac{1}{x}, \qquad \text{if} \qquad x \neq 0, \\ &= 0, \qquad\quad\;\; \text{if} \qquad x = 0, \end{aligned}$$

in the vicinity of the origin is one such example. The reader may be inclined to disregard the seriousness of the situation since this function misbehaves only at one point of the interval, namely, where it fails to have a derivative, but it was demonstrated by Weierstrass that there are continuous functions that fail to have a derivative at any point of the interval. Such pathological behavior of continuous functions led to a careful inquiry into the meaning of such geometrical concepts as the area under a curve, the length of arc, the volume, etc., on which a considerable portion of analysis rested until the latter part of the last century. An offshoot of this inquiry was the theory of functions of a real variable, which is characterized by a complete absence of intuitive geometrical concepts. The exact meaning of the concept of the area under a curve is thus made to depend on an analytical definition of the definite integral, rather than the other way round.

A brief outline of the analytical definition of the Riemann definite integral, together with the statement of some important theorems, is contained in this section. The treatment is necessarily condensed and glosses over some of the more intricate points, which properly must be deferred to a course in the theory of functions of a real variable.

If the elements of a given set of numbers (E), when represented by points on the number axis, possess the property that there are no points of the set (E) to the right of some fixed point, then the set is called *bounded on the right*, or *bounded above*. If, on the other hand, there are no points of the set (E) to the left of some fixed point, then the set (E) is *bounded below*, or *bounded on the left*. For example, the set of positive rational numbers is bounded on the left but is unbounded above. The set of negative rational numbers is bounded above, but not below. A set of real numbers between 0 and 1 is bounded above and below. Every set bounded above and below is said simply to be *bounded*.

If a set of numbers (E) has the following properties:

(*i*) *there is no number of the set* (E) *which is greater than some number* M;

(*ii*) *there is at least one number of the set* (E) *which is greater than* $M - \epsilon$, *where* ϵ *is an arbitrarily small positive number, then the number* M *is called the upper bound of the set* (E).

On the other hand, if the set (E) possesses the properties:

(*i*) *that there is no number of the set smaller than some number* m;

(*ii*) *that there is at least one number of the set* (E) *which is less than* $m + \epsilon$, *however small the positive number* ϵ *may be, then the number* m *is called the lower bound of the set* (E).

If one cares to take advantage of the suggestiveness of the language of geometry, then the meaning of the upper and lower bounds of a set of numbers is the following. The points of the bounded set (E) all fall in a segment of finite length, and the numbers m and M are the end points of the segment. It appears from this remark that every bounded set necessarily must have the upper and lower bounds. A rigorous arithmetical proof of this fact is based on the study of the linear continua and is not given here.*

Denote the set of values assumed by the bounded function $f(x)$ in (a, b) by (E). Then the upper and lower bounds of the set (E) are called the *upper and lower bounds of* $f(x)$ *in* (a, b). Let M and m be the upper and lower bounds of a bounded function $f(x)$ defined in the interval (a, b), and let (x_1, x_2) be any subinterval contained in (a, b). Then the upper bound of $f(x)$ in (x_1, x_2) certainly will not be greater than M, and the lower bound will not be smaller than m. Hence, if the interval (a, b) is subdivided into n subintervals, and if the upper and the lower bounds of $f(x)$ in each of the subintervals are denoted by M_i and m_i, respectively $(i = 1, 2, \cdots, n)$, one is assured that $m_i \geq m$ and $M_i \leq M$. This fact is of basic importance in what follows.

Let the interval (a, b) be divided into n subintervals by the points

$$a \equiv x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n \equiv b,$$

and form the sums

$$S = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \equiv \sum_{i=1}^{n} M_i \, \Delta x_i$$

---

* See, for example, E. W. Hobson, Theory of Functions of a Real Variable

and

$$s = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \equiv \sum_{i=1}^{n} m_i \, \Delta x_i.$$

The sums S and s depend on a particular mode of subdivision of (a, b) into n parts, but for each particular mode of subdivision $s \leq S$, since $m_i \leq M_i$.

Consider the set of numbers S corresponding to all conceivable modes of subdivision. No matter what mode of subdivision is chosen,

$$S = \sum_{i=1}^{n} M_i \, \Delta x_i \leq M(b - a), \tag{35-1}$$

since $M_i \leq M$ and $\sum_{i=1}^{n} \Delta x_i = b - a$.

Similarly,

$$s = \sum_{i=1}^{n} m_i \, \Delta x_i \geq m(b - a). \tag{35-2}$$

Now, since

$$s \leq S,$$

it follows from (35-1) and (35-2) that

$$m(b - a) \leq S \leq M(b - a),$$

and

$$m(b - a) \leq s \leq M(b - a).$$

These inequalities state that for any mode of subdivision of the interval (a, b) into n parts the sums S and s lie between $m(b - a)$ and $M(b - a)$, and therefore the sums S have the lower bound, say J, and the sums s have the upper bound, say I. From the definition of these bounds it follows that

$$J \leq S$$

and

$$I \geq s.$$

If each of the intervals Δx_i (in any particular subdivision) is subdivided further by inserting additional points and the new sum S is formed for this larger number of subdivisions, then this new sum cannot be greater than the old S, and in general it will be less. Moreover, the sum S arising from any mode of subdivision of (a, b) will not be less than the sum s resulting from this or any other mode of subdivision. This leads one to suspect that when the number n of subdivisions is increased indefinitely in such a way that each $\Delta x_i \to 0$, the sum S tends to a definite limit which is independent of the mode of subdividing (a, b) into subintervals Δx_i, and that this limit is J.

On the other hand, s will in general increase with the increase in the number of subdivisions, while remaining less than or equal to S, so that it is likely that its limit is I. The correctness of these surmises was established for the first time in 1875 by the French mathematician Darboux, who enunciated the following theorem:

Theorem. *The sums S and s tend to definite limits J and I, respectively, when the number of intervals Δx_i increases indefinitely in such a way that each and every $\Delta x_i \to 0$. Moreover, $I \leq J$.*

The proof of this theorem is given in books on the theory of functions of a real variable.* The numbers J and I are called, respectively, the *upper* and *lower Riemann integrals of $f(x)$ in (a, b).*

From the definition of the upper and lower bounds it follows that if ξ_i is any value of x in the interval Δx_i, then

$$m_i \leq f(\xi_i) \leq M_i.$$

Therefore,

$$\sum_{i=1}^{n} m_i \, \Delta x_i \leq \sum_{i=1}^{n} f(\xi_i) \, \Delta x_i \leq \sum_{i=1}^{n} M_i \, \Delta x_i,$$

and if it should happen that

$$\lim_{n \to \infty} \sum_{i=1}^{n} m_i \, \Delta x_i = \lim_{n \to \infty} \sum_{i=1}^{n} M_i \, \Delta x_i,$$

that is, if $I = J$, then the sum

$$\sum_{i=1}^{n} f(\xi_i) \, \Delta x_i$$

---

* See Hobson, E. W., The Theory of Functions of a Real Variable, vol. 1, 3d ed., p. 462; de la Vallée Poussin, C. J., Cours d'analyse infinitésimale, vol. 1, 3d ed., p. 250.

will have the same limit for any choice of the points ξ_i. Any bounded function whose upper and lower integrals are equal is said to be *integrable in the sense of Riemann.*

It will be observed that if $I = J$, then the difference between the sums S and s,

$$\sum_{i=1}^{n} (M_i - m_i)\,\Delta x_i,$$

must approach zero as n increases indefinitely and the $\Delta x_i \to 0$. This criterion enables one to show that every continuous function is necessarily integrable. For, if $f(x)$ is continuous in the closed interval (a, b), it is uniformly continuous there. Hence, the interval (a, b) can be divided into a finite number of subintervals Δx_i such that for any pair of values of x (say x_1 and x_2) in the subinterval Δx_i

$$|f(x_1) - f(x_2)| \leq |M_i - m_i| < \epsilon.$$

Thus,

$$\sum_{i=1}^{n} (M_i - m_i)\,\Delta x_i < \sum_{i=1}^{n} \epsilon\,\Delta x_i = \epsilon(b - a).$$

Since $\epsilon(b - a)$ can be made arbitrarily small, any continuous function of x is integrable.

It is not difficult to see that any function which is continuous except for a finite number of ordinary discontinuities is integrable also. The points of discontinuity can be enclosed in a finite number of intervals, each of which is arbitrarily small, and it follows from the discussion just above (and from the fact that the number of intervals enclosing the points of discontinuity is finite) that

$$\sum_{i=1}^{n} (M_i - m_i)\,\Delta x_i$$

can be made arbitrarily small.

If $f(x)$ is integrable then $|f(x)|$ is integrable, since, for any given mode of subdivision, the difference $S - s$ for $|f(x)|$ is less than or equal to the corresponding difference $S - s$ for $f(x)$. Moreover, if $f_1(x) \geq f_2(x)$, and if both f_1 and f_2 are integrable in (a, b), then

$$\int_a^b f_1(x)\,dx \geq \int_a^b f_2(x)\,dx, \qquad \text{if} \qquad b > a.$$

Note that the difference

$$\varphi(x) \equiv f_1(x) - f_2(x) \geq 0,$$

and that $\varphi(x)$ is integrable since both $f_1(x)$ and $f_2(x)$ are. Then

$$\int_a^b \varphi(x)\,dx \geq 0.$$

Thus

$$\int_a^b f_1(x)\,dx - \int_a^b f_2(x)\,dx \geq 0,$$

which proves the assertion.

It may be remarked, in conclusion, that more general definitions of the integral have been given by Lebesgue, Stieltjes, Denjoy, and others. The object of all these definitions is to include a wider range of functions than those integrable in the sense of Riemann. One characteristic feature of all these definitions is that if a function satisfies the limitations imposed upon it in the treatment of Riemann, then the other integrals lead to the same value as the integral of Riemann.

Since only Riemannian integrals are used in this book, the term *integrable function* will be used henceforth to mean *function integrable in the sense of Riemann.* It should be observed that the integrable function, as defined above, is necessarily bounded. Some writers include in the class of integrable functions those unbounded functions whose *improper integrals* exist. The subject of Improper Integrals is treated in Chap. X.

36. Direct Evaluation of Integrals. The definition of the Riemann integral of $f(x)$ can be applied to calculate the value of the integral of any integrable function, but it will be seen from the examples given below that the problem of evaluating an integral directly from the definition is likely to be quite vexing. In fact, one rarely uses mathematical definitions for purposes of calculation,* and a powerful theorem known as the *fundamental theorem of the integral calculus* (established in Sec. 38) enables one to calculate easily the integrals of a large number of continuous functions. However, the concept of the definite integral as the limit of the sum is so important that it seems desirable to apply the definition to the calculation of the integrals of some simple functions before the fundamental theorem is discussed. The reason for dignifying this theorem with the adjective *fundamental* will be apparent when the same problems are solved with its aid.

---

* The reader will recall that he seldom uses the definition of the derivative in calculating the derivatives of specific functions. Instead, he develops a set of formulas based on the definition.

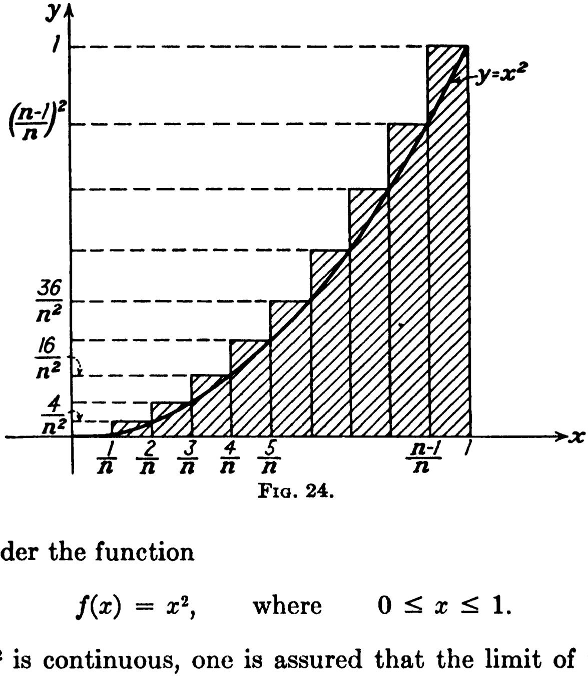

Fig. 24.

Consider the function

$$f(x) = x^2, \qquad \text{where} \qquad 0 \le x \le 1.$$

Since x^2 is continuous, one is assured that the limit of the sum $\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$ exists and is independent of the choice of the ξ_i and of the mode of subdivision of the interval (0, 1) into n parts. Thus, to make the problem easier, let the interval (0, 1) be divided into n equal parts so that $\Delta x_i = \frac{1}{n}$ (Fig. 24). The points of subdivision $x_0, x_1, x_2, \cdots, x_{n-1}, x_n$ are

$$x_0 = 0,\ x_1 = \frac{1}{n},\ x_2 = \frac{2}{n},\ \cdots,\ x_{n-1} = \frac{n-1}{n},\ x_n = 1.$$

If the points ξ_i be chosen as the right end points of the intervals $\Delta x_i = x_i - x_{i-1}$, there results

$$\begin{aligned}\int_0^1 x^2\,dx &= \lim_{n\to\infty}\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i \\ &= \lim_{n\to\infty}\left[\frac{1}{n^2}\cdot\frac{1}{n}+\left(\frac{2}{n}\right)^2\cdot\frac{1}{n}+\cdots+\left(\frac{n-1}{n}\right)^2\frac{1}{n}+1^2\cdot\frac{1}{n}\right] \\ &= \lim_{n\to\infty}\frac{1^2+2^2+3^2+\cdots+n^2}{n^3}.\end{aligned}$$

Now the sum of the terms in the numerator is given by the formula*

$$1^2+2^2+3^2+\cdots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

Substituting this in the right-hand member of the foregoing expression gives

$$\begin{aligned}\int_0^1 x^2\,dx &= \lim_{n\to\infty}\frac{n(n+1)(2n+1)}{6n^3} \\ &= \frac{1}{6}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right) \\ &= \frac{1}{3}.\end{aligned}$$

The reader will find it instructive to carry out the calculation in this example by choosing $\Delta x_i = \frac{1}{n}$ and $\xi_i = (i - \theta_i)\,\Delta x_i$, where $0 \leq \theta_i \leq 1$, $(i = 1, 2, \cdots, n)$. The θ_i need not be equal. Of course, the limit should be the same for an arbitrary choice of the points ξ_i.

Another example of a direct calculation of the integral may prove instructive. Let it be required to evaluate the integral

$$\int_0^x e^x\,dx.$$

If the interval $(0, x)$ is divided into n equal parts, so that

$$\Delta x_i = x_i - x_{i-1} = \frac{x}{n} \equiv h,$$

---

* The reader will have no difficulty in proving this assertion by mathematical induction.

and if ξ_i is chosen to be equal to x_{i-1} (Fig. 25), then

$$s_n = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$$
$$= h[1 + e^h + e^{2h} + \cdots + e^{(n-1)h}].$$

Calculating the sum of the n terms of the geometrical progression appearing in the bracket gives

$$s_n = h\frac{e^{nh} - 1}{e^h - 1} = (e^x - 1)\cdot\frac{h}{e^h - 1},$$

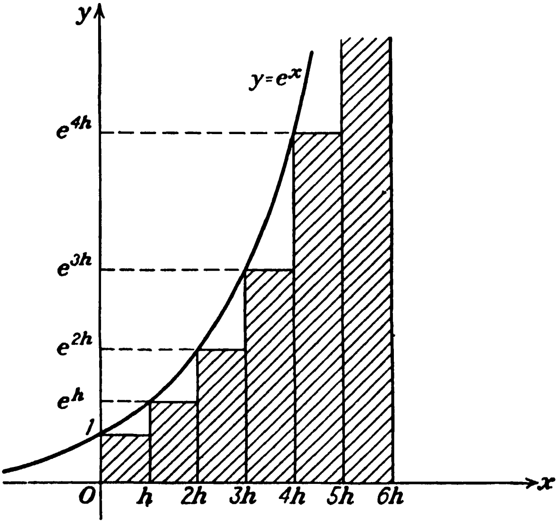

Fig. 25.

since $nh = x$. Letting $h \to 0$ and noting that

$$\lim_{h\to 0}\frac{h}{e^h - 1} = 1$$

gives

$$\int_0^x e^x\,dx = e^x - 1.$$

PROBLEM

Evaluate directly from the definition $\int_2^4 x\,dx$.

37. Mean-value Theorems for Integrals.

Theorem 1 (First Mean-value Theorem for Integrals). *Let $f(x)$ and $\varphi(x)$ be two functions which are continuous in the interval*

(a, b) and suppose that $\varphi(x)$ does not change sign in this interval. Then there exists at least one value ξ, where $a \leq \xi \leq b$, such that

$$\int_a^b f(x)\varphi(x)\,dx = f(\xi)\int_a^b \varphi(x)\,dx.$$

In order to prove this theorem, consider

$$\int_a^b [M - f(x)]\varphi(x)\,dx,$$

in which M stands for the maximum value of $f(x)$ in (a, b). Let it be supposed that $\varphi(x) \geq 0$; then, since $M \geq f(x)$, it follows that*

$$\int_a^b [M - f(x)]\varphi(x)\,dx \geq 0,$$

or that

$$M\int_a^b \varphi(x)\,dx \geq \int_a^b f(x)\varphi(x)\,dx. \tag{37-1}$$

Similarly, if m represents the minimum value of $f(x)$ in (a, b), it is seen that

$$\int_a^b f(x)\varphi(x)\,dx \geq m\int_a^b \varphi(x)\,dx. \tag{37-2}$$

It follows from (37-1) and (37-2) that

$$\int_a^b f(x)\varphi(x)\,dx = \mu\int_a^b \varphi(x)\,dx,$$

in which $m \leq \mu \leq M$. But $f(x)$ is continuous and hence there exists some value ξ between a and b, for which $f(\xi) = \mu$. Therefore

$$\int_a^b f(x)\varphi(x)\,dx = f(\xi)\int_a^b \varphi(x)\,dx. \tag{37-3}$$

In case $\varphi(x) = 1$, (37-3) reduces to

$$\int_a^b f(x)\,dx = f(\xi)(b - a),$$

a form which is already familiar from elementary calculus.

This first mean-value theorem expresses the integral in terms of the value of one of the functions at a point intermediate between

---

* If $\varphi(x) \leq 0$, it is merely necessary to reverse the inequalities.

a and b. Another mean-value theorem for integrals uses the intermediate value in the limits. This second theorem requires the definition of functions which are monotone increasing or monotone decreasing. *A function $\varphi(x)$ is said to be monotone increasing in the interval (a, b), if $\varphi(x_2) \geq \varphi(x_1)$ for $x_2 > x_1$, for all values of x_1 and x_2 in (a, b). It is said to be monotone decreasing, if $\varphi(x_2) \leq \varphi(x_1)$ for $x_2 > x_1$, for all values of x_1 and x_2 in (a, b).*

Theorem 2 (Second Mean-value Theorem for Integrals). *Let $f(x)$ and $\varphi(x)$ be two functions which are continuous in (a, b). If $\varphi(x)$ is a positive monotone-decreasing function, then there exists a value ξ, where $a \leq \xi \leq b$, such that*

$$\int_a^b f(x)\varphi(x)\,dx = \varphi(a)\int_a^\xi f(x)\,dx.$$

If $\varphi(x)$ is a positive monotone-increasing function in (a, b), then there exists a value ξ, where $a \leq \xi \leq b$, such that

$$\int_a^b f(x)\varphi(x)\,dx = \varphi(b)\int_\xi^b f(x)\,dx.$$

If $\varphi(x)$ is either monotone increasing or monotone decreasing but not necessarily always positive, then

$$\int_a^b f(x)\varphi(x)\,dx = \varphi(a)\int_a^\xi f(x)\,dx + \varphi(b)\int_\xi^b f(x)\,dx.$$

The proof of the first part of the theorem depends upon the fact that

$$\int_a^b f(x)\varphi(x)\,dx \equiv \lim_{n\to\infty}\sum_{i=1}^n f(\xi_i)\varphi(\xi_i)(x_i - x_{i-1}), \qquad x_{i-1} \leq \xi_i \leq x_i,$$

in which the interval (a, b) has been divided into n parts by the points x_i in such a way that

$$a \equiv x_0 < x_1 < x_2 \cdots < x_{n-1} < x_n \equiv b.$$

Now consider the expressions

$$s_r = \sum_{i=1}^r f(\xi_i)(x_i - x_{i-1}), \qquad r = 1, 2, \cdots, n,$$

and let M represent the largest and m represent the smallest of these expressions. Then, since

$$f(\xi_1)(x_1 - x_0) = s_1$$

and

$$f(\xi_j)(x_j - x_{j-1}) = s_j - s_{j-1}, \qquad j = 2, 3, \cdots, n,$$

it follows that

$$\begin{aligned}\sum_{i=1}^{n}\varphi(\xi_i)f(\xi_i)(x_i - x_{i-1}) &= \varphi(\xi_1)s_1 + \sum_{i=2}^{n}\varphi(\xi_i)(s_i - s_{i-1})\\ &= \sum_{i=1}^{n-1}s_i[\varphi(\xi_i) - \varphi(\xi_{i+1})] + \varphi(\xi_n)s_n. \qquad (37\text{-}4)\end{aligned}$$

But $\varphi(x)$ is a positive monotone-decreasing function, so that each expression $\varphi(\xi_i) - \varphi(\xi_{i+1})$ is nonnegative, as is $\varphi(\xi_n)$. Hence the right-hand member of (37-4) is not less than

$$\sum_{i=1}^{n-1}m[\varphi(\xi_i) - \varphi(\xi_{i+1})] + m\varphi(\xi_n) = m\varphi(\xi_1),$$

and it is not greater than

$$\sum_{i=1}^{n-1}M[\varphi(\xi_i) - \varphi(\xi_{i+1})] + M\varphi(\xi_n) = M\varphi(\xi_1).$$

Therefore,

$$m\varphi(\xi_1) \leqq \sum_{i=1}^{n}\varphi(\xi_i)f(\xi_i)(x_i - x_{i-1}) \leq M\varphi(\xi_1).$$

As n is increased so that each of the differences $x_i - x_{i-1}$ approaches zero, ξ_1 approaches a. Moreover, since m and M are the smallest and greatest of the expressions

$$s_r = \sum_{i=1}^{r}f(\xi_i)(x_i - x_{i-1}),$$

it is seen that, as n is increased, m and M approach, respectively, the minimum and maximum values $\bar{m}$ and $\bar{M}$ of $\int_a^\eta f(x)\,dx$, where η takes all values from a to b. Hence,

$$\bar{m}\varphi(a) \leq \int_a^b \varphi(x)f(x)\,dx \leq \bar{M}\varphi(a)$$

and

$$\int_a^b \varphi(x)f(x)\,dx = \mu\varphi(a),$$

in which $\bar{m} \leq \mu \leq \bar{M}$. But $f(x)$ is continuous and therefore* there must be some value ξ for which $\mu = \int_a^\xi f(x)\,dx$. It follows that

---

* See Sec. 38.

$$(37\text{-}5) \qquad \int_a^b \varphi(x)f(x)\,dx = \varphi(a)\int_a^\xi f(x)\,dx, \qquad a \le \xi \le b.$$

Similarly, if $\varphi(x)$ is a positive monotone-increasing function, it can be shown that

$$(37\text{-}6) \qquad \int_a^b \varphi(x)f(x)\,dx = \varphi(b)\int_\xi^b f(x)\,dx, \qquad a \le \xi \le b.$$

For the third part of the proof, suppose that $\varphi(x)$ is a monotone-increasing function. Then define $\Phi(x)$ by

$$\Phi(x) \equiv \varphi(b) - \varphi(x)$$

so that $\Phi(x)$ is a positive monotone-decreasing function in (a, b). Hence,

$$\begin{aligned}\int_a^b f(x)\Phi(x)\,dx &= \Phi(a)\int_a^\xi f(x)\,dx\\ &= [\varphi(b) - \varphi(a)]\int_a^\xi f(x)\,dx.\end{aligned}$$

But

$$\begin{aligned}\int_a^b f(x)\Phi(x)\,dx &= \int_a^b f(x)[\varphi(b) - \varphi(x)]\,dx\\ &= \varphi(b)\int_a^b f(x)\,dx - \int_a^b f(x)\varphi(x)\,dx.\end{aligned}$$

Hence,

$$\int_a^b f(x)\varphi(x)\,dx = \varphi(b)\int_a^b f(x)\,dx - \varphi(b)\int_a^\xi f(x)\,dx + \varphi(a)\int_a^\xi f(x)\,dx,$$

or

$$\int_a^b f(x)\varphi(x)\,dx = \varphi(b)\int_\xi^b f(x)\,dx + \varphi(a)\int_a^\xi f(x)\,dx.$$

The same result can be obtained in the case where $\varphi(x)$ is a monotone-decreasing function by the use of the auxiliary function

$$\Psi(x) = \varphi(x) - \varphi(b).$$

As an example of the application of the first mean-value theorem for integrals, consider the problem of establishing the bounds for the elliptic integral

$$\int_0^{\frac{1}{2}} \frac{dx}{\sqrt{(1 - x^2)(1 - k^2x^2)}}, \qquad \text{where} \qquad k^2 < 1.$$

It follows from the foregoing that

$$\int_0^{1/2} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \frac{1}{\sqrt{1-k^2\xi^2}} \int_0^{1/2} \frac{dx}{\sqrt{1-x^2}},$$

where $0 \leq \xi \leq \frac{1}{2}$. But

$$\int_0^{1/2} \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x \Big|_0^{1/2} = \frac{\pi}{6}.$$

Hence,

$$\int_0^{1/2} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \frac{\pi}{6} \frac{1}{\sqrt{1-k^2\xi^2}}.$$

Replacing ξ by zero and one-half gives a lower and an upper bound for the value of the elliptic integral. Thus,

$$\frac{\pi}{6} \leq \int_0^{1/2} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \leq \frac{\pi}{6} \frac{1}{\sqrt{1-\frac{1}{4}k^2}}.$$

38. Fundamental Theorem of the Integral Calculus. Consider an integrable function $f(x)$ defined in the interval $a \leq x \leq b$. It is clear that if x is any point of the interval, then

$$\int_a^x f(x)\, dx$$

exists, and its magnitude depends on the value assigned to the upper limit. Thus, for a fixed lower limit, the definite integral is a function of the upper limit, and it will be designated by $F(x)$, so that

$$F(x) = \int_a^x f(x)\, dx.$$

In order to avoid confusion between the letter x denoting the variable of integration and the same letter used to denote the upper limit, the integral will be written as

$$F(x) = \int_a^x f(t)\, dt.$$

Let $x + h$ be some point of the interval (a, b), then

$$F(x + h) = \int_a^{x+h} f(t)\, dt,$$

and

$$F(x+h) - F(x) = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt$$
$$= \int_a^{x+h} f(t)\,dt + \int_x^a f(t)\,dt.$$

Combining the integrals in the right-hand member of this expression gives

$$(38\text{-}1) \qquad F(x+h) - F(x) = \int_x^{x+h} f(t)\,dt.$$

An important deduction can be drawn immediately from (38-1). Since

$$\lim_{h\to 0} \int_x^{x+h} f(t)\,dt = 0,$$

it follows that

$$\lim_{h\to 0} [F(x+h) - F(x)] = 0$$

so that $F(x)$ *is a continuous function whenever* $f(x)$ *is an integrable function.* Furthermore, if $f(x)$ is a continuous function in the closed interval (a, b), then the application to (38-1) of the first mean-value theorem for integrals gives

$$F(x+h) - F(x) = f(\xi)h,$$

where ξ lies between x and $x + h$.
Then

$$(38\text{-}2) \qquad \frac{dF}{dx} = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} f(\xi) = f(x).$$

Now any function $F(x)$ whose derivative is equal to $f(x)$ is called *a primitive* of $f(x)$, or *an indefinite integral* of $f(x)$. Thus the statement embodied in (38-2) gives an important theorem.

Theorem. *The definite integral, regarded as a function of the upper limit, is a primitive of the integrand whenever the latter is a continuous function.*

If $G(x)$ is any primitive of $f(x)$, then it differs from $F(x)$ at most by a constant, since the derivatives of $G(x)$ and $F(x)$ are equal.*

---

* See Sec. 20, p. 53.

Thus

$$G(x) = F(x) + C,$$

or

$$G(x) = \int_a^x f(x)\,dx + C.$$

Setting the upper limit of the integral equal to a gives

$$G(a) = C,$$

since $\int_a^a f(x)\,dx = 0.$
Consequently,

$$\int_a^x f(x)\,dx = G(x) - G(a), \tag{38-3}$$

where $G(x)$ is any primitive.

Equation (38-3) is the symbolic statement of the following theorem:

Fundamental Theorem of Integral Calculus. *If $f(x)$ is continuous in the interval $a \leq x \leq b$ and $G(x)$ is a function such that $\frac{dG}{dx} = f(x)$ for all values of x in this interval, then*

$$\int_a^b f(x)\,dx = G(b) - G(a).$$

In the first example of Sec. 36 it was shown directly from the definition of the definite integral as the limit of the sum that

$$\int_0^1 x^2\,dx = \tfrac{1}{3}.$$

Since a primitive of x^2 is $\frac{x^3}{3}$, it follows from the fundamental theorem that

$$\int_0^1 x^2\,dx = \frac{x^3}{3}\bigg|_0^1 = \frac{1}{3}.$$

Again

$$\int_0^x e^x\,dx = e^x\Big|_0^x = e^x - 1,$$

which agrees with the result obtained in Sec. 36 by a different method.

The symbol $\int f(x)\,dx$ is used to denote an indefinite integral of $f(x)$. The reader will deduce the following properties of indefinite integrals:

(a) $d\int f(x)\,dx = f(x)\,dx$;

(b) $\int cf(x)\,dx = c\int f(x)\,dx$, where c is a constant;

(c) $\int [f_1(x) + f_2(x)]\,dx = \int f_1(x)\,dx + \int f_2(x)\,dx$.

39. Differentiation under the Integral Sign. Consider the definite integral

$$\int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\,dx, \tag{39-1}$$

in which the integrand is a function of a parameter α (as well as of x) and the limits are also functions of α. This integral is a function of α, and not of the variable of integration x. It is frequently desired to obtain the derivative of (39-1) with respect to α in cases in which it is inconvenient or impossible to express the indefinite integral in explicit form. For this reason the following theorem is of great value.

Theorem. *Let*

$$\varphi(\alpha) = \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\,dx,$$

where u_0 *and* u_1 *are differentiable functions in a closed interval* (α_0, α_1); $f(x, \alpha)$ *and* $f_\alpha(x, \alpha)$ *are continuous in the region* $\alpha_0 \leq \alpha \leq \alpha_1$, $u_0(\alpha) \leq x \leq u_2(\alpha)$. *Then*

$$\frac{d\varphi}{d\alpha} = \int_{u_0(\alpha)}^{u_1(\alpha)} \frac{\partial f(x, \alpha)}{\partial \alpha}\,dx - f(u_0, \alpha)\frac{du_0}{d\alpha} + f(u_1, \alpha)\frac{du_1}{d\alpha}.$$

In order to prove this theorem, first form

$$\Delta\varphi \equiv \varphi(\alpha + \Delta\alpha) - \varphi(\alpha).$$

Then

$$\begin{aligned}
\Delta\varphi &= \int_{u_0(\alpha+\Delta\alpha)}^{u_1(\alpha+\Delta\alpha)} f(x, \alpha + \Delta\alpha)\,dx - \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\,dx \\
&= \int_{u_0(\alpha+\Delta\alpha)}^{u_0(\alpha)} f(x, \alpha + \Delta\alpha)\,dx + \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha + \Delta\alpha)\,dx \\
&\quad + \int_{u_1(\alpha)}^{u_1(\alpha+\Delta\alpha)} f(x, \alpha + \Delta\alpha)\,dx - \int_{u_0(\alpha)}^{u_1(\alpha)} f(x, \alpha)\,dx
\end{aligned}$$

The first integral of the first expression has been written as the sum of the first three integrals of the second expression. If the second and fourth integrals are combined, one can write

$$\frac{\Delta\varphi}{\Delta\alpha} = \int_{u_0(\alpha)}^{u_1(\alpha)} \frac{f(x, \alpha + \Delta\alpha) - f(x, \alpha)}{\Delta\alpha}\,dx - \int_{u_0(\alpha)}^{u_0(\alpha+\Delta\alpha)} \frac{f(x, \alpha + \Delta\alpha)}{\Delta\alpha}\,dx + \int_{u_1(\alpha)}^{u_1(\alpha+\Delta\alpha)} \frac{f(x, \alpha + \Delta\alpha)}{\Delta\alpha}\,dx.$$

By the mean-value theorem of the differential calculus*

$$f(x, \alpha + \Delta\alpha) - f(x, \alpha) = \Delta\alpha \cdot f_\alpha(x, \eta), \qquad \alpha < \eta < \alpha + \Delta\alpha.$$

Also, by the first mean-value theorem for integrals,

$$\int_{u_0(\alpha)}^{u_0(\alpha+\Delta\alpha)} \frac{f(x, \alpha + \Delta\alpha)}{\Delta\alpha}\,dx = \frac{f(\xi_0, \alpha + \Delta\alpha)}{\Delta\alpha}[u_0(\alpha + \Delta\alpha) - u_0(\alpha)]$$
$$= f(\xi_0, \alpha + \Delta\alpha)\left[\frac{u_0(\alpha + \Delta\alpha) - u_0(\alpha)}{\Delta\alpha}\right],$$

where $u_0(\alpha) \leq \xi_0 \leq u_0(\alpha + \Delta\alpha)$; and, similarly,

$$\int_{u_1(\alpha)}^{u_1(\alpha+\Delta\alpha)} \frac{f(x, \alpha + \Delta\alpha)}{\Delta\alpha}\,dx = f(\xi_1, \alpha + \Delta\alpha)\left[\frac{u_1(\alpha + \Delta\alpha) - u_1(\alpha)}{\Delta\alpha}\right],$$

where $u_1(\alpha) \leq \xi_1 \leq u_1(\alpha + \Delta\alpha)$.
Therefore,

$$\frac{\Delta\varphi}{\Delta\alpha} = \int_{u_0(\alpha)}^{u_1(\alpha)} \frac{\partial f(x, \eta)}{\partial\alpha}\,dx - f(\xi_0, \alpha + \Delta\alpha)\frac{\Delta u_0}{\Delta\alpha} + f(\xi_1, \alpha + \Delta\alpha)\frac{\Delta u_1}{\Delta\alpha}$$

with

$$\alpha < \eta < \alpha + \Delta\alpha, \qquad u_0(\alpha) \leq \xi_0 \leq u_0(\alpha + \Delta\alpha),$$
$$u_1(\alpha) \leq \xi_1 \leq u_1(\alpha + \Delta\alpha).$$

Hence

$$\text{(39-2)} \quad \frac{d\varphi}{d\alpha} = \int_{u_0(\alpha)}^{u_1(\alpha)} \frac{\partial f(x, \alpha)}{\partial\alpha}\,dx - f(u_0, \alpha)\frac{du_0}{d\alpha} + f(u_1, \alpha)\frac{du_1}{d\alpha}.$$

In case the limits $u_0(\alpha)$ and $u_1(\alpha)$ are constants, say x_0 and x_1, (39-2) reduces to

$$\text{(39-3)} \qquad \frac{d\varphi}{d\alpha} \equiv \frac{d}{d\alpha}\left[\int_{x_0}^{x_1} f(x, \alpha)\,dx\right] = \int_{x_0}^{x_1} \frac{\partial f(x, \alpha)}{\partial\alpha}\,dx.$$

---

* See Sec. 20.

Example 1. Formula (39-3) is frequently used for evaluating definite integrals.* Thus, if

$$\varphi(\alpha) = \int_0^\pi \log\,(1 + \alpha \cos x)\,dx,$$

then

$$\begin{aligned}\varphi'(\alpha) &= \int_0^\pi \frac{\cos x}{1 + \alpha\cos x}\,dx = \frac{1}{\alpha}\int_0^\pi \left(1 - \frac{1}{1 + \alpha\cos x}\right) dx \\ &= \frac{1}{\alpha}\left(x + \frac{1}{\sqrt{1-\alpha^2}}\sin^{-1}\frac{\alpha + \cos x}{1 + \alpha\cos x}\right)\Bigg|_0^\pi \\ &= \frac{1}{\alpha}\left[\pi + \frac{1}{\sqrt{1-\alpha^2}}\left(\sin^{-1}\frac{\alpha - 1}{1 - \alpha} - \sin^{-1}\frac{\alpha + 1}{1 + \alpha}\right)\right] \\ &= \frac{1}{\alpha}\left(\pi + \frac{-\pi}{\sqrt{1-\alpha^2}}\right) = \frac{\pi}{\alpha} - \frac{\pi}{\alpha\sqrt{1-\alpha^2}}.\end{aligned}$$

Therefore,

$$\begin{aligned}\varphi(\alpha) &= \pi\int\left(\frac{1}{\alpha} - \frac{1}{\alpha\sqrt{1-\alpha^2}}\right) d\alpha \\ &= \pi\left(\log\alpha + \log\frac{1 + \sqrt{1-\alpha^2}}{\alpha}\right) + c\end{aligned}$$

or

$$\varphi(\alpha) = \pi\log\,(1 + \sqrt{1-\alpha^2}) + c.$$

But, when $\alpha = 0$,

$$\varphi(0) = \int_0^\pi \log 1\,dx = 0.$$

Hence,

$$0 = \pi\log 2 + c \qquad \text{and} \qquad c = -\pi\log 2,$$

and

$$\varphi(\alpha) = \pi\log\left(\frac{1 + \sqrt{1-\alpha^2}}{2}\right)$$

Example 2. Find $\dfrac{d\varphi}{d\alpha}$, if $\varphi(\alpha) = \displaystyle\int_{-\alpha^2}^{2\alpha} e^{-\frac{x^2}{\alpha^2}}\,dx.$

Then

$$\begin{aligned}\frac{d\varphi}{d\alpha} &= \int_{-\alpha^2}^{2\alpha} \frac{2x^2}{\alpha^3} e^{-\frac{x^2}{\alpha^2}}\,dx - e^{-\alpha^2}(-2\alpha) + e^{-4}(2) \\ &= \int_{-\alpha^2}^{2\alpha} \frac{2x^2}{\alpha^3} e^{-\frac{x^2}{\alpha^2}}\,dx + 2\alpha e^{-\alpha^2} + 2e^{-4}.\end{aligned}$$

---

* See also Sec. 97.

PROBLEMS

1. Find $\dfrac{d\varphi}{d\alpha}$, if $\varphi(\alpha) = \displaystyle\int_0^{\pi} (1 - \alpha \cos x)^2\, dx$.

2. Find $\dfrac{d\varphi}{d\alpha}$, if $\varphi(\alpha) = \displaystyle\int_0^{\alpha^2} \tan^{-1} \frac{x}{\alpha^2}\, dx$.

3. Find $\dfrac{d\varphi}{d\alpha}$, if $\varphi(\alpha) = \displaystyle\int_0^{\alpha} \tan (x - \alpha)\, dx$.

4. Find $\dfrac{d\varphi}{dx}$, if $\varphi(x) = \displaystyle\int_0^{x^2} \sqrt{x}\, dx$.

5. Differentiate under the sign and thus evaluate

$$\varphi(\alpha) = \int_0^1 \frac{x^\alpha - 1}{\log x}\, dx.$$

What is the permissible range of values of α?

6. Differentiate under the sign and thus evaluate

$\displaystyle\int_0^{\pi} (\alpha - \cos x)^{-2}\, dx$ by using $\displaystyle\int_0^{\pi} (\alpha - \cos x)^{-1}\, dx = \pi(\alpha^2 - 1)^{-1}$, $\alpha^2 > 1$.

7. Show that

$$\begin{aligned} \int_0^{\pi} \log (1 - 2\alpha \cos x + \alpha^2)\, dx &= 0, && \text{if} \quad \alpha^2 \leq 1, \\ &= \pi \log \alpha^2, && \text{if} \quad \alpha^2 \geq 1. \end{aligned}$$

8. Verify that

$$y = \frac{1}{k}\int_0^x f(\alpha) \sin k(x - \alpha)\, d\alpha$$

is a solution of the differential equation

$$\frac{d^2y}{dx^2} + k^2y = f(x),$$

where k is a constant.

40. Change of Variable. In performing the evaluation of the definite integral with the aid of the fundamental theorem, it is frequently desirable to change the variable of integration in the integral $\int_a^b f(x)\, dx$ by means of some relation $x = \varphi(t)$. In order to emphasize the need of caution, observe the following example of a purely formal change of variable which leads to a nonsensical result.

The integral

$$\int_{-1}^{1} x^2\,dx = \frac{x^3}{3}\bigg|_{-1}^{1} = \frac{2}{3}.$$

Now, let it be supposed that the variable x is replaced by another variable t connected with x by means of the relation

$$t = x^2.$$

The limit in the resulting integral, corresponding to $x = -1$, is

$$t = (-1)^2 = 1$$

and the upper limit becomes

$$t = (1)^2 = 1.$$

Thus, in the transformed integral the upper and lower limits are each equal to unity, so that the value of the transformed integral appears to be zero. The difficulty here lies in the fact that the function $x = \varphi(t)$ is not a single-valued function.

Theorem. *Let $f(x)$ be a continuous and single-valued function of x in the closed interval (a, b) and let $x = \varphi(t)$ be a continuous, single-valued function of t possessing a continuous derivative in the closed interval (α, β), where $\varphi(\alpha) = a$ and $\varphi(\beta) = b$. Then*

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f[\varphi(t)]\varphi'(t)\,dt.$$

In order to prove this theorem, consider the integrals

$$(40\text{-}1)\qquad F(x) = \int_a^x f(x)\,dx \qquad \text{and} \qquad \Phi(t) = \int_\alpha^t f[\varphi(t)]\varphi'(t)\,dt.$$

By hypothesis $x = \varphi(t)$, so that $F(x)$ is a function of t whose derivative with respect to t is

$$\frac{dF(x)}{dt} = \frac{dF(x)}{dx}\frac{dx}{dt}.$$

But $\dfrac{dF(x)}{dx} = f(x)$, since $f(x)$ is continuous, so that

$$\frac{dF}{dt} = f(x)\varphi'(t) = f[\varphi(t)]\varphi'(t), \qquad \text{where} \qquad \varphi'(t) = \frac{dx}{dt}.$$

On the other hand, the derivative of $\Phi(t)$ is

$$\frac{d\Phi(t)}{dt} = f[\varphi(t)]\varphi'(t),$$

since the integrand of Φ is continuous.*

The derivatives of F and Φ with respect to t being equal, the functions F and Φ can differ only by a constant. But for $t = \alpha$, $x = \varphi(\alpha) = a$, so that

$$F(x)_{t=\alpha} = F(a) = 0.$$

Moreover,

$$\Phi(\alpha) = 0.$$

Thus the functions $F(x)$ and $\Phi(t)$ are equal for $t = \alpha$, and consequently, they are equal for all values of t in the interval (α, β). Setting $t = \beta$ in the second of the integrals (40-1) and

$$x = \varphi(\beta) = b$$

in the first integral gives the desired result

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f[\varphi(t)]\varphi'(t)\,dt.$$

Example. Consider $\int_0^1 \sqrt{1 - x^2}\,dx$. If $x = \sin t$, then

$$dx = \cos t\,dt.$$

When $x = 0$, $t = 0$ and when $x = 1$, $t = \frac{\pi}{2}$, so that

$$\int_0^1 \sqrt{1 - x^2}\,dx = \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 t}\cos t\,dt$$

$$= \int_0^{\frac{\pi}{2}} \cos^2 t\,dt = \frac{\pi}{4}.$$

PROBLEM

Justify the substitution $dx = \varphi'(t)\,dt$ in the indefinite integral $\int f(x)\,dx$ to obtain the indefinite integral $\int f[\varphi(t)]\varphi'(t)\,dt$.

41. Applications of Definite Integrals. The reader is familiar with the derivation of the formulas for areas bounded by curves, for volumes and surfaces of revolution, and for the length of arc of a curve. These formulas are given below for reference, and in the light of the meaning of the definite integral stated in Sec. 35, it is best to regard them as the definitions of the geometrical entities under consideration.

---

* See Theorems 1 and 2, Sec. 12.

The outline of the procedure employed in defining the length of arc of a curve is as follows. Assume that the equation of a continuous curve C is given in a parametric form as

$$\begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t), \qquad a \leqq t \leq b. \end{cases} \tag{41-1}$$

Let it be required to find the length of arc of C between the points corresponding to the values of the parameter $t = a$ and $t = b$. Divide the interval $a \leq t \leq b$ into n subintervals of lengths $\Delta t_1, \Delta t_2, \cdots, \Delta t_n$ by inserting $n - 1$ points of subdivision chosen so that

$$a \equiv t_0 < t_1 < t_2 \cdots < t_n \equiv b.$$

Denote the increments of the functions in (41-1), corresponding to $\Delta t_1, \Delta t_2, \cdots, \Delta t_n$, by $\Delta x_1, \Delta x_2, \cdots, \Delta x_n$, and $\Delta y_1, \Delta y_2, \cdots, \Delta y_n$.

The sum of the lengths of the chords ΔC_i joining the consecutive points on the curve C which correspond to the values of the parameter $t = t_i$, $(i = 1, 2, \cdots, n)$, is

$$s_n = \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2},$$

where

$$\Delta C_i \equiv \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}.$$

If the limit of the sum s_n exists* as the number of subdivisions is increased indefinitely in such a way that each $\Delta t_i \to 0$, then this limit is called the length of arc of the curve between the points corresponding to the values of the parameter $t = a$ and $t = b$.

To ensure the existence of this limit, it is not sufficient to require merely that $\varphi_1(t)$ and $\varphi_2(t)$ be continuous. It is sufficient, however, to assume the continuity of the derivatives of $\varphi_1(t)$

---

* Of course, this limit must be independent of the mode of subdivision.

and $\varphi_2(t)$ in the interval (a, b). In this event the length s of the curve is given by the formula

$$s = \int^{b} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

The volume of the solid of revolution which is generated by revolving a continuous curve $y = f(x)$ about the x-axis, and which is bounded by the planes $x = x_1$ and $x = x_2$ can be calculated from

$$V = \int_{x_1}^{x_2} \pi y^2\, dx.$$

The surface of revolution of the same solid is

$$\begin{aligned} S &= \int_{x_1}^{x_2} 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= \int_{y_1}^{y_2} 2\pi y \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy. \end{aligned}$$

In the case of the surface of revolution it is sufficient to demand the continuity of the derivative of $y = f(x)$, so that the integrand will be a continuous function.

Example 1. Consider the problem of determining the length of arc of the logarithmic spiral whose equation is

$$\rho = ae^{n\theta}.$$

The element of arc length in polar coordinates is given by

$$ds = \sqrt{(d\rho)^2 + (\rho\, d\theta)^2},$$

so that

$$\begin{aligned} s &= a\int_0^{\theta} e^{n\theta}\sqrt{n^2 + 1}\, d\theta \\ &= a\frac{\sqrt{n^2 + 1}}{n}(e^{n\theta} - 1). \end{aligned}$$

Example 2. Find the length of the lemniscate $\rho^2 = a^2 \cos 2\theta$. Now $\rho = a\sqrt{\cos 2\theta}$, and therefore,

$$\frac{d\rho}{d\theta} = -\frac{a \sin 2\theta}{\sqrt{\cos 2\theta}}.$$

Hence,

$$s = 4a\int_0^{\frac{\pi}{4}} \sqrt{\frac{\sin^2 2\theta}{\cos 2\theta} + \cos 2\theta}\, d\theta$$

$$= 4a\int_0^{\frac{\pi}{4}} \frac{d\theta}{\sqrt{\cos 2\theta}}.$$

Setting $2\theta = x$ gives

$$s = 2a\int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{\cos x}},$$

and the substitution of $\cos \varphi = \sqrt{\cos x}$ gives

$$s = 2\sqrt{2}a\int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - \frac{1}{2}\sin^2 \varphi}}.$$

This integral cannot be evaluated in a closed form with the aid of the elementary functions. It is the standard form of the elliptic integral of the first kind, which can be evaluated either in infinite series (see Sec. 86) or with the aid of the tables of elliptic integrals.

·PROBLEMS

1. Find the area bounded by the curve

$$x^{2/3} + y^{2/3} = a^{2/3}.$$

2. Find the volume generated by revolving the curve $x^{2/3} + y^{2/3} = a^{2/3}$ about the x-axis.

3. Find the surface of the solid whose volume is required in Prob. 2.

4. Find the surface of the ellipsoid of revolution

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1.$$

Ans. $S = \frac{2\pi ac}{e}\sin^{-1} e + 2\pi ac\sqrt{1 - e^2}$, if $c > a$;

$S = \frac{\pi a^2(1 - e^2)}{e}\log\frac{1 + e}{1 - e} + 2\pi a^2$, if $c < a$;

where e is the eccentricity.

5. Find the surface generated by revolving the cardioid

$$\rho = 2a\cos^2\frac{\theta}{2}$$

about the polar axis. *Ans.* $^{32}/_{5}\,\pi a^2$.

CHAPTER V

MULTIPLE INTEGRALS

42. Double Integrals. The definition of the double integral is entirely analogous to that given earlier for the simple integral.

Let $f(x, y)$ be a single-valued bounded function in a closed region R (Fig. 26). Let the region R be subdivided in any manner into n subregions $\Delta R_1, \Delta R_2, \cdots, \Delta R_n$ of areas $\Delta A_1, \Delta A_2, \cdots, \Delta A_n$. The upper and lower bounds of the function $f(x, y)$ in the subregion ΔR_i will be denoted by M_i and m_i, respectively. Then

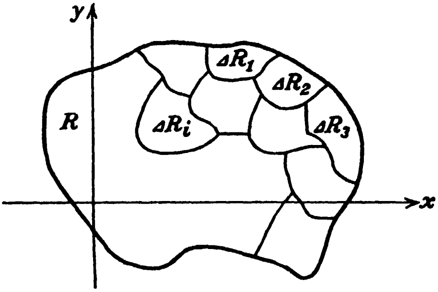

Fig. 26.

$$m_i \, \Delta A_i \leq f(\xi_i, \eta_i) \, \Delta A_i \leq M_i \, \Delta A_i,$$

where (ξ_i, η_i) is any point in the region ΔR_i. Let s_n and S_n denote the sums

$$s_n \equiv \sum_{i=1}^{n} m_i \, \Delta A_i \qquad \text{and} \qquad S_n \equiv \sum_{i=1}^{n} M_i \, \Delta A_i,$$

which clearly depend on the number n of subregions and on the mode of subdividing the region R into parts.

If the sums s_n and S_n approach the same limit S when the number of subregions is increased indefinitely in such a way that each $\Delta A_i \to 0$, then the limit of the sum

$$s'_n \equiv \sum_{i=1}^{n} f(\xi_i, \eta_i) \, \Delta A_i$$

will also be S, since

$$\sum_{i=1}^{n} m_i \, \Delta A_i \leq \sum_{i=1}^{n} f(\xi_i, \eta_i) \, \Delta A_i \leq \sum_{i=1}^{n} M_i \, \Delta A_i,$$

for any mode of subdivision and for any choice of the points (ξ_i, η_i). This common limit S is defined as the *double integral* of $f(x, y)$ over the region R, and one writes

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(\xi_i, \eta_i) \, \Delta A_i \equiv \int_R f(x, y) \, dA. \tag{42-1}$$

The region R is called the *region of integration* and corresponds to the interval of integration (a, b) in the case of the simple integral. The term *double integral* refers to the dimensionality of the region R.

The existence of the unique limit S is guaranteed if one assumes that the function $f(x, y)$ is continuous in the closed region R. This is obvious from the geometrical considerations of Sec. 44 if one assumes that the function $z = f(x, y)$ is represented by a surface. The analytical proof makes use of the property of the uniform continuity of continuous functions $f(x, y)$ defined over a closed region R and is not given here.*

PROBLEM

Let M and m be the maximum and the minimum values, respectively, of the continuous single-valued function $f(x, y)$ defined over a region R; then

$$\sum_{i=1}^{n} m \, \Delta A_i \leq \sum_{i=1}^{n} f(\xi_i, \eta_i) \, \Delta A_i \leq \sum_{i=1}^{n} M \, \Delta A_i.$$

Show, from a consideration of this inequality, that

$$\int_R f(x, y) \, dA = f(\xi, \eta) A,$$

where (ξ, η) is some point of the region R, whose area is A. This is the mean-value theorem for double integrals.

43. Evaluation of the Double Integral. It will be assumed in this section that the function $f(x, y)$ is continuous over a closed region R, so that the limit of the sum (42-1) will be independent of the mode of dividing the region R into the subregions ΔR_i and also of the choice of the point (ξ_i, η_i) within the subregion ΔR_i.

---

* See the analogous discussion in Sec. 35.

The definition given in the preceding section gives no convenient means for the evaluation of double integrals. In order to evaluate the double integral, it will be simpler to consider first

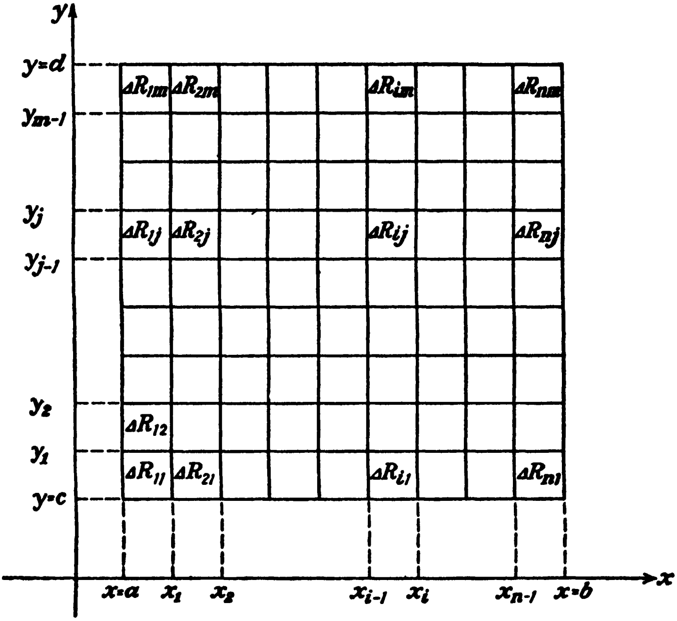

FIG. 27.

the case in which the region R (Fig. 27) is a rectangle bounded by the lines $x = a$, $x = b$, $y = c$, $y = d$. The extension to other types of regions will be indicated later. Subdivide R into mn rectangles by drawing the lines $x = x_1$, $x = x_2$, $\cdots$, $x = x_{n-1}$, $y = y_1$, $y = y_2$, $\cdots$, $y = y_{m-1}$. Define

$$\Delta x_i \equiv x_i - x_{i-1},$$

where $x_0 = a$ and $x_n = b$, and define

$$\Delta y_j \equiv y_j - y_{j-1},$$

where $y_0 = c$ and $y_m = d$. Let ΔR_{ij} be the rectangle bounded by the lines $x = x_{i-1}$, $x = x_i$, $y = y_{j-1}$, $y = y_j$. Denote the area of the rectangle ΔR_{ij}, which stands in the ith column and the jth row, by

$$\Delta A_{ij} = \Delta x_i \, \Delta y_j$$

and form the sum of the expressions

$$f(\xi_{ij}, \eta_{ij}) \, \Delta A_{ij},$$

where (ξ_{ij}, η_{ij}) is any point of ΔR_{ij}. This sum is

$$\sum_{i=1,\, j=1}^{i=n,\, j=m} f(\xi_{ij}, \eta_{ij}) \, \Delta x_i \, \Delta y_j, \tag{43-1}$$

where the summation symbol $\sum\limits_{i=1,\, j=1}^{i=n,\, j=m}$ signifies that the terms appearing under the summation sign can be added in any manner whatsoever.

Suppose that the terms of (43-1) are arranged so that the rectangles ΔR_{i1} are used first; then all the rectangles ΔR_{i2}; then all the rectangles ΔR_{i3}; etc. This is equivalent to taking the sum of the terms corresponding to the first, second, third, etc., rows of rectangles and then adding these sums. Such a rearrangement of the terms of the sum permits one to write (43-1) as the double sum

$$\sum_{j=1}^{m} \Delta y_j \left[\sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij}) \, \Delta x_i \right], \tag{43-2}$$

where the expression in the bracket represents the sum of the terms corresponding to any fixed value of j. From the definition of the ordinary definite integral, the expression appearing in the bracket approaches a limit when $n \to \infty$ in such a way that each $\Delta x_i \to 0$. Hence, one can write

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij}) \, \Delta x_i = \int_a^b f(x, \eta_j) \, dx,$$

or

$$\sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij}) \, \Delta x_i = \int_a^b f(x, \eta_j) \, dx + \epsilon_j, \tag{43-3}$$

where $\lim\limits_{n \to \infty} \epsilon_j = 0$. The latter expression results from the definition of the limit, and it merely states that for a sufficiently large n the difference between the sum and the integral can be made as small as desired.

Now the integral

$$\int_a^b f(x, \eta_j)\, dx$$

contains η_j as a parameter and hence defines a function of η_j, say

$$\varphi(\eta_j) = \int_a^b f(x, \eta_j)\, dx. \tag{43-4}$$

Substituting from (43-3) in (43-2) and making use of (43-4) gives

$$\sum_{j=1}^{m} \Delta y_j \left[\sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij})\, \Delta x_i \right] = \sum_{j=1}^{m} \Delta y_j [\varphi(\eta_j) + \epsilon_j]. \tag{43-5}$$

But

$$\lim_{m \to \infty} \sum_{j=1}^{m} \varphi(\eta_j)\, \Delta y_j = \int_c^d \varphi(y)\, dy$$

so that

$$\sum_{j=1}^{m} \varphi(\eta_j)\, \Delta y_j = \int_c^d \varphi(y)\, dy + \epsilon,$$

where $\lim_{m \to \infty} \epsilon = 0$. Thus (43-5) becomes

$$\sum_{j=1}^{m} \Delta y_j \left[\sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij})\, \Delta x_i \right] = \int_c^d \varphi(y)\, dy + \sum_{j=1}^{m} \epsilon_j\, \Delta y_j + \epsilon.$$

Calculating the limit as both m and n tend to infinity gives*

---

* The reason for the vanishing of $\lim_{\substack{m \to \infty \\ n \to \infty}} \sum_{j=1}^{m} \epsilon_j\, \Delta y_j$ is not entirely obvious. Even though $\epsilon_j \to 0$ when $n \to \infty$, the index m likewise tends to infinity, and it is conceivable that the limit of the sum may be different from zero. The fact that this is not the case here can be seen from the following considerations. Let $\int_{x_{i-1}}^{x_i} f(x, \eta_j)\, dx$ denote the integral of the function $f(x, \eta_j)$ over the interval Δx_i. Then

$$\left| \int_{x_{i-1}}^{x_i} f(x, \eta_j)\, dx - f(\xi_{ij}, \eta_{ij})\, \Delta x_i \right| \le \left| M_{ij} - m_{ij} \right| \Delta x_i,$$

where M_{ij} and m_{ij} stand for the maximum and minimum values of the function $f(x, \eta_j)$ in the interval (x_{i-1}, x_i). But

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \sum_{j=1}^{m} \Delta y_j\left[\sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij})\,\Delta x_i\right] = \int_c^d \varphi(y)\,dy$$

$$= \int_c^d\left[\int_a^b f(x, y)\,dx\right]dy,$$

where the second step is obtained by recalling the definition (43-4). The left-hand member of this equation gives the double integral $\int_R f(x, y)\,dA$, so that

$$\int_R f(x, y)\,dA = \int_c^d \int_a^b f(x, y)\,dx\,dy. \tag{43-6}$$

The double integral is, therefore, evaluated by considering $f(x, y)$ as a function of x alone but containing y as a parameter,

---

$$\sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x, \eta_j)\,dx = \int_a^b f(x, \eta_j)\,dx.$$

Hence,

$$\left|\int_a^b f(x, \eta_j)\,dx - \sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij})\,\Delta x_i\right| \leq \sum_{i=1}^{n} \left|M_{ij} - m_{ij}\right| \Delta x_i.$$

Let ϵ' be the largest of the numbers $|M_{ij} - m_{ij}|$, $(i = 1, 2, \cdots, n)$, $(j = 1, 2, \cdots, m)$, then

$$\sum_{i=1}^{n} \left|M_{ij} - m_{ij}\right| \Delta x_i \leq \epsilon' \sum_{i=1}^{n} \Delta x_i = \epsilon'(b - a).$$

Thus

$$\left|\int_a^b f(x, \eta_j)\,dx - \sum_{i=1}^{n} f(\xi_{ij}, \eta_{ij})\,\Delta x_i\right| \leq \epsilon'(b - a).$$

But, from (43-3), the left-hand member of this expression is precisely the numerical value of ϵ_j, that is,

$$|\epsilon_j| \leq \epsilon'(b - a).$$

Hence,

$$\left|\sum_{j=1}^{m} \epsilon_j\,\Delta y_j\right| \leq \epsilon'(b - a) \sum_{j=1}^{m} \Delta y_j = \epsilon'(b - a)(c - d).$$

Since $f(x, y)$ is a continuous function, the difference between the maximum and the minimum values of the function in any subregion ΔR_{ij} tends to zero as the number of subregions is increased indefinitely, and consequently, $\epsilon' \to 0$.

and integrating it between $x = a$ and $x = b$ and then integrating the resulting function of y between $y = c$ and $y = d$. The right-hand member of (43-6) is known as an *iterated integral*, and (43-6) establishes the relation between the double integral over the rectangle R and an iterated integral over the same rectangle.

Similarly, by taking the sum of the terms in each column and then adding these sums, there results

$$(43\text{-}7)\qquad \int_R f(x, y)\, dA = \int_a^b \int_c^d f(x, y)\, dy\, dx.$$

In case (43-7) is used, $f(x, y)$ is first considered as a function of y alone and integrated between $y = c$ and $y = d$, and then the resulting function of x is integrated between $x = a$ and $x = b$. Either (43-6) or (43-7) can be used, but one of them is frequently simpler in the case of a particular function $f(x, y)$.

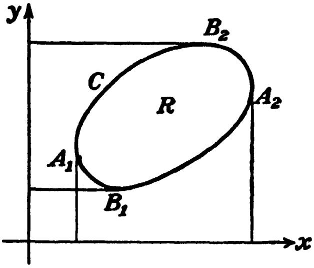

FIG. 28.

Suppose R is not a rectangle but is a region bounded by a closed curve C (Fig. 28) which is cut by any line parallel to either axis in at most two points. Let B_1 and B_2 be the points of C having the minimum and maximum ordinates, and let A_1 and A_2 be the points of C having the minimum and maximum abscissas. Let $x = \varphi_1(y)$ be the equation of $B_1A_1B_2$, and $x = \varphi_2(y)$ be the equation of $B_1A_2B_2$. Then, in taking the sum of the terms by rows and adding these sums, the limits for the first integration will be $\varphi_1(y)$ and $\varphi_2(y)$, instead of the constants a and b. The limits for the second integration will be β_1 and β_2, in which β_1 is the y-coordinate of B_1 and β_2 is the y-coordinate of B_2. Then (43-6) is replaced by

$$(43\text{-}8)\qquad \int_R f(x, y)\, dA = \int_{\beta_1}^{\beta_2} \int_{\varphi_1(y)}^{\varphi_2(y)} f(x, y)\, dx\, dy.$$

Similarly, if $y = f_1(x)$ is the equation of $A_1B_1A_2$, $y = f_2(x)$ is the equation of $A_1B_2A_2$, α_1 is the abscissa of A_1, and α_2 is the abscissa of A_2, (43-7) is replaced by

$$(43\text{-}9)\qquad \int_R f(x, y)\, dA = \int_{\alpha_1}^{\alpha_2} \int_{f_1(x)}^{f_2(x)} f(x, y)\, dy\, dx.$$

In case R is a region bounded by a closed curve C which is cut in more than two points by some parallel to one of the axes, the previous results can be applied to subregions of R whose boundaries satisfy the previous conditions. By adding algebraically the integrals over these subregions, the double integral over R is obtained (Fig. 29).

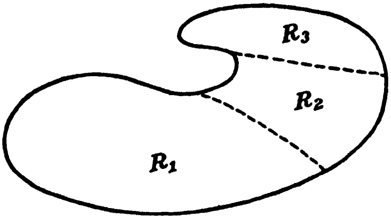

FIG. 29.

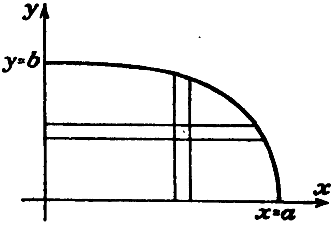

FIG. 30.

Example 1. Compute the value of $I_1 = \int_R y\, dA$ where R is the region in the first quadrant bounded by the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{(Fig. 30).}$$

$$x = \sqrt{a^2 - \frac{a^2}{b^2} y^2}$$

Using (43-8) and summing first by rows, one finds

$$x = \frac{a}{b}\sqrt{b^2-y^2}$$

$$I_1 = \int_0^b \int_0^{\frac{a}{b}\sqrt{b^2-y^2}} y\, dx\, dy = \int_0^b \left(yx \Big|_0^{\frac{a}{b}\sqrt{b^2-y^2}} \right) dy$$
$$= \frac{a}{b}\int_0^b y\sqrt{b^2 - y^2}\, dy = -\frac{a}{3b}(b^2 - y^2)^{\frac{3}{2}}\Big|_0^b$$
$$= \frac{ab^2}{3}.$$

Using (43-9) yields

$$I_1 = \int_0^a \int_0^{\frac{b}{a}\sqrt{a^2-x^2}} y\, dy\, dx = \int_0^a \left(\frac{y^2}{2} \Big|_0^{\frac{b}{a}\sqrt{a^2-x^2}} \right) dx$$
$$= \frac{b^2}{2a^2}\int_0^a (a^2 - x^2)\, dx = \frac{b^2}{2a^2}\left(a^2x - \frac{x^3}{3}\right)\Big|_0^a = \frac{ab^2}{3}.$$

It may be remarked that the value of I_1 is equal to $\bar{y}A$, where $\bar{y}$ is the y-coordinate of the center of gravity of this quadrant of the ellipse, and A is its area. Since $A = \frac{\pi ab}{4}$,

$$\bar{y} = \frac{I_1}{A} = \frac{\dfrac{ab^2}{3}}{\dfrac{\pi ab}{4}} = \frac{4b}{3\pi}.$$

Similarly, by evaluating $I_2 = \int_R x\,dA = \dfrac{a^2b}{3}$,

$$\bar{x} = \frac{I_2}{A} = \frac{\dfrac{a^2b}{3}}{\dfrac{\pi ab}{4}} = \frac{4a}{3\pi},$$

which is the x-coordinate of the center of gravity.

Example 2. Moment of Inertia. It will be recalled that the moment of inertia of a particle about an axis is the product of its mass by the square of its distance from the axis. If it is desired to find the moment of inertia of a plane region about an axis perpendicular to the plane of the region, the method of Sec. 43 can be applied, where $f(x, y)$ is the square of the distance from the point (x, y) of the region to the axis. Then

$$M = \int_R r^2\,dA.$$

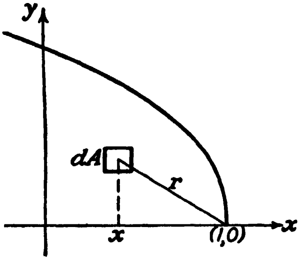

FIG. 31.

For example, let it be required to find the moment of inertia of the area in the first quadrant (Fig. 31), bounded by the parabola $y^2 = 1 - x$ and the coordinate axes, about an axis perpendicular to the xy-plane at $(1, 0)$. The distance from any point $P(x, y)$ to $(1, 0)$ is $r = \sqrt{(x-1)^2 + y^2}$. Therefore,

$$M = \int_R [(x-1)^2 + y^2]\,dA.$$

Evaluating this integral by means of (43-8), there results

$$\begin{aligned} M &= \int_0^1 \int_0^{1-y^2} [(x-1)^2 + y^2]\,dx\,dy \\ &= \int_0^1 \left[\frac{(x-1)^3}{3} + xy^2\right]_0^{1-y^2} dy \\ &= \int_0^1 \left(-\frac{y^6}{3} + y^2 - y^4 + \frac{1}{3}\right) dy = \frac{44}{105}. \end{aligned}$$

44. Geometric Interpretation of the Double Integral. If $f(x, y)$ is a continuous and single-valued function defined over the region R (Fig. 32) of the xy-plane, then $z = f(x, y)$ is the equation of a surface. Let C be the closed curve that is the boundary of R. Using R as a base, construct a cylinder having its elements parallel to the z-axis. This cylinder intersects $z = f(x, y)$ in a curve Γ, whose projection on the xy-plane is C. Denote by S the portion of $z = f(x, y)$ that is enclosed by Γ. Let R be subdivided as in Sec. 43 by the lines $x = x_i$, $(i = 1, 2, \cdots, n - 1)$, and $y = y_j$, $(j = 1, 2, \cdots, m - 1)$. Through each line $x = x_i$ pass a plane parallel to the yz-plane; and through each line $y = y_j$ pass a plane parallel to the xz-plane. The rectangle ΔR_{ij}, whose area is $\Delta A_{ij} = \Delta x_i \, \Delta y_j$, will be the base of a rectangular prism of height $f(\xi_{ij}, \eta_{ij})$, whose volume is approximately equal to the volume enclosed between the surface and the xy-plane by the planes $x = x_{i-1}$, $x = x_i$, $y = y_{j-1}$ and $y = y_j$. Then the sum

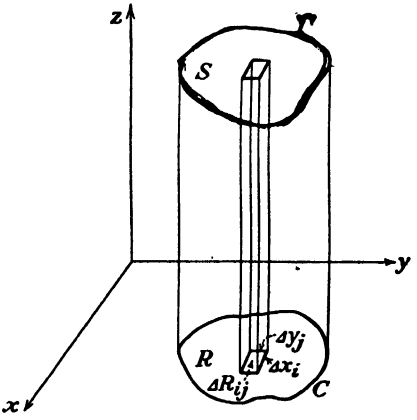

FIG. 32.

(44-1)
$$\sum_{i=1,\, j=1}^{i=n,\, j=m} f(\xi_{ij}, \eta_{ij}) \, \Delta x_i \, \Delta y_j$$

gives an approximate value for the volume V of the portion of the cylinder enclosed between $z = f(x, y)$ and the xy-plane. As $n \to \infty$ and $m \to \infty$, the sum (44-1) approaches V, so that

(44-2)
$$V = \int_R f(x, y) \, dA.$$

The integral in (44-2) can be evaluated by (43-8) in which the prisms are added first in the x-direction, or by (43-9) in which the prisms are added first in the y-direction.

It should be noted that formulas (43-8) and (43-9) give the magnitude of the area of the region R if the function $f(x, y) = 1$; for the left-hand member becomes

$$\int_R dA = \int\int_R dx\,dy,$$

which is A. The area A can be evaluated by using

$$\int_{\beta_1}^{\beta_2}\int_{\varphi_1(y)}^{\varphi_2(y)} dx\,dy \qquad \text{or} \qquad \int_{\alpha_1}^{\alpha_2}\int_{f_1(x)}^{f_2(x)} dy\,dx.$$

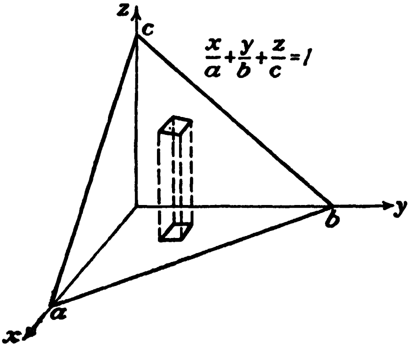

Fig. 33.

Example. Find the volume of the tetrahedron bounded by the plane $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$ and the coordinate planes (Fig. 33). Here

$$z = c\left(1 - \frac{x}{a} - \frac{y}{b}\right).$$

If the prisms are summed first in the x-direction, they will be summed from $x = 0$ to the line ab, whose equation is

$$\frac{x}{a} + \frac{y}{b} = 1.$$

Therefore,

$$V = \int_0^b \int_0^{a\left(1-\frac{y}{b}\right)} c\left(1 - \frac{x}{a} - \frac{y}{b}\right) dx\,dy$$

$$= c\int_0^b \left(x - \frac{x^2}{2a} - \frac{xy}{b}\right)\Bigg|_0^{a\left(1-\frac{y}{b}\right)} dy$$

$$= ac\int_0^b \left(\frac{1}{2} - \frac{y}{b} + \frac{y^2}{2b^2}\right) dy$$

$$= \frac{abc}{6}.$$

This result was obtained by using (43-8) for the evaluation of V, but (43-9) could be used equally well.

PROBLEMS

1. Evaluate

(*a*) $\int_0^1 \int_x^{\sqrt{x}} (x^2 + y^2)\, dy\, dx;$

(*b*) $\int_{-1}^2 \int_{x^2}^{x+2} dy\, dx;$

(*c*) $\int_0^\pi \int_0^{a(1-\cos\theta)} \rho\, d\rho\, d\theta,$

and describe the regions of integration in (*a*) and (*b*).

2. Verify that $\int_R (x^2 + y^2)\, dy\, dx = \int_R (x^2 + y^2)\, dx\, dy,$ where the region R is a triangle formed by the lines $y = 0$, $y = x$, and $x = 1$.

3. Evaluate and describe the regions of integration for

(*a*) $\int_0^a \int_{a-x}^{\sqrt{a^2-x^2}} y\, dy\, dx;$

(*b*) $\int_0^a \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dy\, dx;$

(*c*) $\int_0^1 \int_x^{\sqrt{x}} (1 + x^2 + y^2)\, dy\, dx;$

(*d*) $\int_{-1}^2 \int_x^{x+2} dy\, dx;$

(*e*) $\int_1^2 \int_2^5 xy\, dx\, dy.$

4. Find the areas enclosed by the following pairs of curves:

(*a*) $y = x$, $y = x^2$;
(*b*) $y = 2 - x$, $y^2 = 2(2 - x)$;
(*c*) $y = 4 - x^2$, $y = 4 - 2x$;
(*d*) $y^2 = 5 - x$, $y = x + 1$;
(*e*) $y = \sqrt{a^2 - x^2}$, $y = a - x$.

5. Find by double integration the volume of one of the wedges cut off from the cylinder $x^2 + y^2 = a^2$ by the planes $z = 0$ and $z = x$.

6. Find the volume of the solid bounded by the paraboloid $y^2 + z^2 = 4x$ and the plane $x = 5$.

7. Find the volume of the solid bounded by the plane $z = 0$, the surface $z = x^2 + y^2 + 2$, and the cylinder $x^2 + y^2 = 4$.

8. Find the smaller of the areas bounded by $y = 2 - x$ and $x^2 + y^2 = 4$.

9. Find the volume bounded by the cylinders $y = x^2$, $y^2 = x$, and the planes $z = 0$ and $z = 1$.

10. Find the volume of the solid bounded by the cylinders $x^2 + y^2 = a^2$ and $y^2 + z^2 = a^2$.

11. Find the coordinates of the center of gravity of the area enclosed by $y = 4 - x^2$ and $y = 4 - 2x$.

12. Find the moments of inertia about the x- and y-axes of the smaller of the areas enclosed by $y = a - x$ and $x^2 + y^2 = a^2$.

45. Triple Integrals. The triple integral is defined in a manner entirely analogous to the definition of the double integral. The function

$$f(x, y, z)$$

is to be continuous and single-valued over the region of space R enclosed by the surface S. Let R be subdivided into subregions ΔR_{ijk}. If ΔV_{ijk} is the volume of ΔR_{ijk}, the triple integral of $f(x, y, z)$ over R is defined by

$$(45\text{-}1) \qquad \int_R f(x, y, z)\, dV \equiv \lim_{n, m, p \to \infty} \sum_{i=1, j=1, k=1}^{i=n, j=m, k=p} f(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})\, \Delta V_{ijk}$$

in a way entirely analogous to that given in Sec. 43.

In order to evaluate the triple integral, R is considered to be subdivided by planes parallel to the three coordinate planes and the case of the rectangular parallelepiped is treated first. In this case

$$\Delta V_{ijk} = \Delta x_i\, \Delta y_j\, \Delta z_k.$$

By suitably arranging the terms of the sum

$$\sum_{i=1, j=1, k=1}^{i=n, j=m, k=p} f(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})\, \Delta x_i\, \Delta y_j\, \Delta z_k,$$

it can be shown, as in Sec. 43, that

$$(45\text{-}2) \qquad \int_R f(x, y, z)\, dV = \int_{z_0}^{z_1} \int_{y_0}^{y_1} \int_{x_0}^{x_1} f(x, y, z)\, dx\, dy\, dz.$$

By other arrangements of the terms of the sum, the triple integral can be expressed by means of iterated integrals in which the order of integration is any permutation of that given in (45-2).

If R is not a rectangular parallelepiped, the triple integral over R will be evaluated by iterated integrals in which the limits for the first two integrations will be functions instead of constants. By extending the method of Sec. 43, it is readily shown that

$$(45\text{-}3)\qquad \int_R f(x, y, z)\, dV = \int_{z_1}^{z_2}\int_{f_1(z)}^{f_2(z)}\int_{\varphi_1(y,z)}^{\varphi_2(y,z)} f(x, y, z)\, dx\, dy\, dz.$$

Similarly the triple integral can be evaluated by interchanging the order of integration in the iterated integral and suitably choosing the limits.

The expression (45-3), or the similar expressions obtained by a different choice of the order of integration, gives the formula for the volume of R in case $f(x, y, z) \equiv 1$. Therefore,

$$V = \int_{z_1}^{z_2}\int_{f_1(z)}^{f_2(z)}\int_{\varphi_1(y,z)}^{\varphi_2(y,z)} dx\, dy\, dz.$$

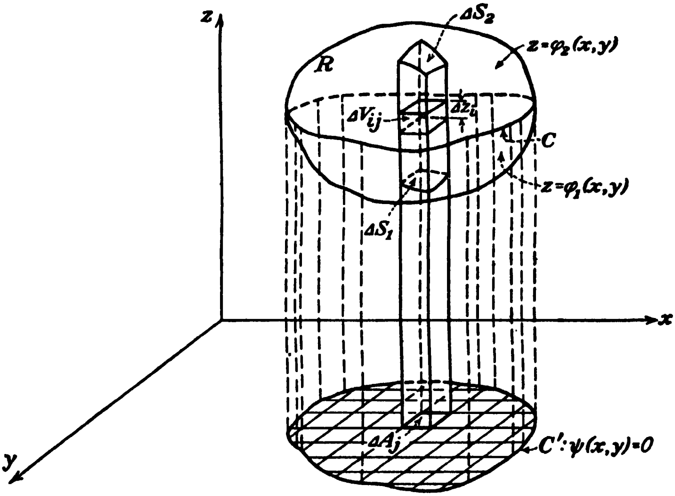

FIG. 34.

In order to clarify the method of evaluating a triple integral by means of iterated integrals, consider the calculation of the integral of the function $f(x, y, z)$ over a region R, which is bounded by a closed surface S. Let $\Phi(x, y, z) = 0$ be the equation of S and denote by V the volume enclosed by S (Fig. 34). The surface S will be assumed to be such that a straight line parallel to one of the coordinate axes, say z, cuts S in not more than two points. An extension to more complicated surfaces can be made immediately in a manner analogous to that indicated in Sec. 43 for double integrals.

Construct the right cylinder which projects the surface S on the xy-plane. Let C be the curve along which the projecting cylinder is tangent to S, and let the projection of C on the xy-plane be the curve C' whose equation is $\psi(x, y) = 0$. The curve C divides the surface S into two parts. Denote the equation of the upper part of the surface by $z = \varphi_2(x, y)$, and that of the lower part by $z = \varphi_1(x, y)$.* Next, divide the volume V into elementary volume elements ΔV in the following way. Divide the area enclosed by C', by a series of lines parallel to the coordinate axes, into a number of elementary areas, of which ΔA_j is typical. Using each of the elements ΔA_j as a base, erect on each element a prism that will cut S in ΔS_1 and ΔS_2, and divide the portions of the prisms lying between ΔS_1 and ΔS_2 by a series of planes parallel to the xy-plane. The entire volume V is thus divided into elementary volumes ΔV. The element $\Delta V_{ij} = \Delta z_i \, \Delta A_j$ is outlined in Fig. 34.

Now select a point (x, y, z) within each of the volume elements ΔV_{ij}, and form the sum

$$\sum_i [f(x, y, z) \, \Delta z_i] \, \Delta A_j.$$

The limit of this sum as $\Delta z_i \to 0$ defines the integral

$$\left[\int_{z=\varphi_1(x,y)}^{z=\varphi_2(x,y)} f(x, y, z) \, dz\right] \Delta A_j. \tag{45-4}$$

The quantity within the brackets is a function of x and y, say $F(x, y)$, and the limit of the sum

$$\sum_j F(x, y) \, \Delta A_j$$

as $\Delta A_j \to 0$ is a double integral over the area bounded by C. If it is evaluated by the methods of Sec. 43, there results

$$\int_R f(x, y, z) \, dV = \int_a^b \int_{y=\psi_1(x)}^{y=\psi_2(x)} \left[\int_{z=\varphi_1(x,y)}^{z=\varphi_2(x,y)} f(x, y, z) \, dz\right] dy \, dx.$$

---

* Obviously $z = \varphi_2(x, y)$ and $z = \varphi_1(x, y)$ are the solutions of the equation $\Phi(x, y, z) = 0$ for z in terms of x and y.

A specific illustration may clarify this matter further. Let it be required to determine the volume enclosed by the ellipsoid

$$\frac{(x-2)^2}{1}+\frac{(y-9)^2}{4}+\frac{(z-16)^2}{9}=1.$$

The equations $z = \varphi_1(x, y)$ and $z = \varphi_2(x, y)$ are, in this case,

$$z = 16 \pm 3\sqrt{1-(x-2)^2-\tfrac{1}{4}(y-9)^2},$$

the upper sign giving φ_2 and the lower giving φ_1. The equation of the curve C' is that of the ellipse

$$\frac{(x-2)^2}{1}+\frac{(y-9)^2}{4}=1,$$

so that

$$\psi_1 = 9 - 2\sqrt{1-(x-2)^2}$$

and

$$\psi_2 = 9 + 2\sqrt{1-(x-2)^2}.$$

The evaluation of the integrals is left as an exercise for the reader.

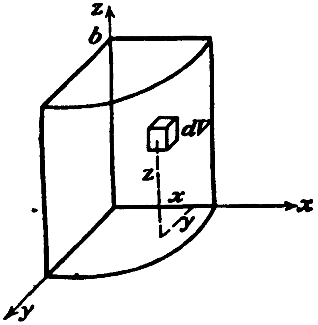

FIG. 35.

As another illustration, let it be required to find the moment of inertia I_x of the solid bounded by the cylinder $x^2 + y^2 = a^2$ and the planes $z = 0$ and $z = b$ about the x-axis (Fig. 35). Assume uniform density σ. The function $f(x, y, z)$ is the square of the distance of any point $P(x, y, z)$ from the x-axis. Therefore,

$$f(x, y, z) = y^2 + z^2.$$

Hence,

$$\begin{aligned}
I_x &= \int_R (y^2+z^2)\sigma\,dV \\
&= 4\sigma\int_0^a\int_0^{\sqrt{a^2-x^2}}\int_0^b (y^2+z^2)\,dz\,dy\,dx \\
&= 4\sigma\int_0^a\int_0^{\sqrt{a^2-x^2}}\left(by^2+\frac{b^3}{3}\right)dy\,dx \\
&= 4\sigma\int_0^a\left(\frac{by^3}{3}+\frac{b^3y}{3}\right)\Bigg|_0^{\sqrt{a^2-x^2}}dx \\
&= \frac{4\sigma b}{3}\int_0^a (a^2+b^2-x^2)\sqrt{a^2-x^2}\,dx
\end{aligned}$$

$$= \frac{4\sigma a^2 b}{3}\int_0^{\frac{\pi}{2}} (a^2 + b^2 - a^2 \sin^2 \theta) \cos^2 \theta \, d\theta$$

$$= \frac{4\sigma a^2 b}{3}\left[(a^2 + b^2)\frac{\pi}{4} - \frac{a^2\pi}{16}\right] = \frac{\sigma a^2 b\pi}{12}(3a^2 + 4b^2).$$

PROBLEMS

1. Evaluate:

(*a*) $\int_0^2\int_1^3\int_1^2 xy^2z \, dz \, dy \, dx$;

(*b*) $\int_0^a\int_0^{\sqrt{a^2-y^2}}\int_0^{\sqrt{a^2-x^2}} dz \, dx \, dy$

(*c*) $\int_0^2\int_0^{\sqrt{2x-x^2}}\int_0^{\frac{x^2+y^2}{4}} dz \, dy \, dx.$

2. Find by triple integration:

(*a*) The volume in the first octant bounded by the coordinate planes and the plane $x + 2y + 3z = 4$.

(*b*) The volume of one of the wedges cut off from the cylinder $x^2 + y^2 = a^2$ by the planes $z = 0$ and $z = x$. *Ans.* $\frac{2}{3}a^3$.

(*c*) The volume enclosed by the cylinder $x^2 + y^2 = 1$ and the planes $z = 0$ and $z = 2 - x$.

(*d*) The volume enclosed by the cylinders $y^2 = z$ and $x^2 + y^2 = a^2$, and by the plane $z = 0$. *Ans.* $\frac{\pi a^4}{4}$.

(*e*) The volume enclosed by the cylinders $y^2 + z^2 = a^2$ and $x^2 + y^2 = a^2$. *Ans.* $\frac{16}{3}a^3$.

(*f*) The volume enclosed by $y^2 + 2z^2 = 4x - 8$, $y^2 + z^2 = 4$, and $x = 0$. *Ans.* 11π.

(*g*) The volume in the first octant bounded by the coordinate planes and $x + 3y + 2z = 6$.

(*h*) The volume enclosed by the cylinder $x^2 + y^2 = 9$ and the planes $z = 5 - x$ and $z = 0$.

(*i*) The volume of the cap cut off from $y^2 + z^2 = 4x$ by the plane $z = x$.

3. Find the moments of inertia about the coordinate axes of the solids in Prob. 2.

4. Find the coordinates of the center of gravity of each of the volumes in Prob. 2.

5. Find by triple integration the moment of inertia of the volume of a hemisphere about a diameter.

6. Find the coordinates of the center of gravity of the volume of the solid in Prob. 5.

7. Find by triple integration the moment of inertia of the volume of the cone $y^2 + z^2 = a^2x^2$ about its axis.

8. Find the moment of inertia of the cone in Prob. 7 about a diameter of its base.

9. Find the volume in the first octant bounded by $z = x + 1$, $x = 0$, $y = 0$, $x = 2z$, and $x^2 + y^2 = 4$.

10. Find the coordinates of the center of gravity of the volume bounded by $z = 2(2 - x - y)$, $z = 0$, and $z = 4 - x^2 - y^2$.

46. Change of Variables in a Double Integral. Let the variables x and y be connected with some other variables u and v by means of the relations

$$\begin{cases} u = f_1(x, y), \\ v = f_2(x, y), \end{cases} \tag{46-1}$$

where the functions entering in (46-1) are continuous and possess continuous first partial derivatives with respect to x and y in some region of the xy-plane. Moreover, assume that the equations (46-1) can be solved for x and y in terms of u and v to yield

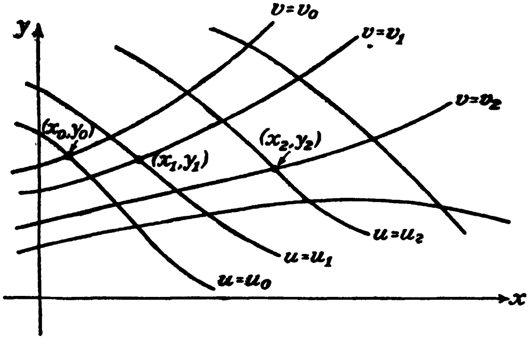

FIG. 36.

$$\begin{cases} x = \varphi_1(u, v), \\ y = \varphi_2(u, v). \end{cases} \tag{46-2}$$

If u and v are assigned some fixed values, say u_0 and v_0, the equations

$$u_0 = f_1(x, y),$$
$$v_0 = f_2(x, y),$$

determine two curves which will intersect in a point (x_0, y_0), such that

$$x_0 = \varphi_1(u_0, v_0),$$
$$y_0 = \varphi_2(u_0, v_0).$$

Thus the pair of numbers (u_0, v_0) determines the point (x_0, y_0), in the xy-plane (Fig. 36).

If u and v are assigned a sequence of constant values

$$(u_1, v_1), (u_2, v_2), (u_3, v_3), \cdots, (u_n, v_n), \cdots,$$

there will be determined a network of curves that will intersect in the points

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \cdots, (x_n, y_n), \cdots.$$

Corresponding to any point whose rectangular coordinates are (x, y) there will be a pair of curves $u =$ const. and $v =$ const., which pass through this point. The totality of numbers (u, v) defines a curvilinear coordinate system, and the curves themselves are called the *coordinate lines*.

Thus, if

$$u = \sqrt{x^2 + y^2},$$
$$v = \tan^{-1}\frac{y}{x},$$

the family of curves $u =$ const. is a family of circles, whereas $v =$ const. defines a family of radial lines. The curvilinear coordinate system, in this case, is the ordinary polar coordinate system (Fig. 37).

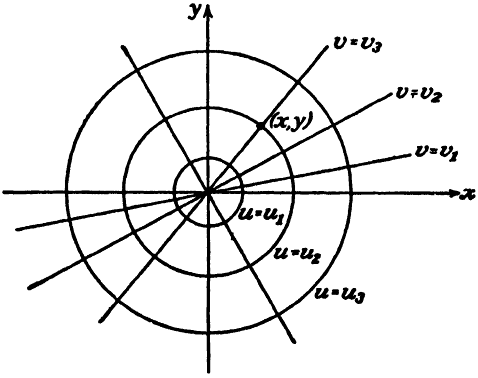

FIG. 37.

Consider the element of area dA in the curvilinear coordinate system (u, v) (Fig. 38) bounded by a quadrilateral $P_1P_2P_4P_3$, the boundary of which is formed by the curves

$$u = f_1(x, y), \qquad u + du = f_1(x, y);$$
$$v = f_2(x, y), \qquad v + dv = f_2(x, y).$$

The positive quantities du and dv are assumed to be finite but may be chosen as small as desired.

The rectangular coordinates (x_1, y_1) of the point P_1 can be calculated from (46-2). Thus,

$$x_1 = \varphi_1(u, v),$$
$$y_1 = \varphi_2(u, v).$$

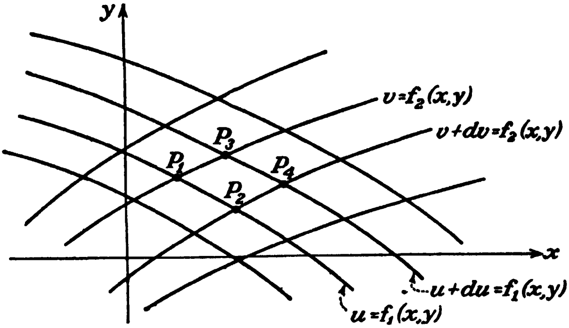

FIG. 38.

Along the boundary P_1P_2, the coordinate u does not change, so that the coordinates (x_2, y_2) of P_2 are given by

$$x_2 = \varphi_1(u, v + dv),$$
$$y_2 = \varphi_2(u, v + dv).$$

Similarly, the coordinates of P_3 and P_4 are

$$x_3 = \varphi_1(u + du, v), \qquad y_3 = \varphi_2(u + du, v);$$
$$x_4 = \varphi_1(u + du, v + dv), \qquad y_4 = \varphi_2(u + du, v + dv).$$

It follows from the theorem of the mean that

$$x_2 = \varphi_1(u, v + dv) = \varphi_1(u, v) + \frac{\partial \varphi_1}{\partial v}\,dv,$$

$$y_2 = \varphi_2(u, v + dv) = \varphi_2(u, v) + \frac{\partial \varphi_2}{\partial v}\,dv,$$

$$x_3 = \varphi_1(u + du, v) = \varphi_1(u, v) + \frac{\partial \varphi_1}{\partial u}\,du,$$

$$y_3 = \varphi_2(u + du, v) = \varphi_2(u, v) + \frac{\partial \varphi_2}{\partial u}\,du.$$

Now the area dA of the curvilinear quadrilateral $P_1P_2P_4P_3$ is approximately equal to twice the area of the rectilinear triangle $P_1P_2P_3$, and the approximation can be made as close as desired by taking du and dv sufficiently small. The area of the rectilinear triangle $P_1P_2P_3$ is given by the determinant*

$$\pm \tfrac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix},$$

where the plus or minus sign is chosen so as to make the area positive.

Substituting the coordinates of the points P_1, P_2, and P_3 in the determinant gives

$$\pm\frac{1}{2}\begin{vmatrix} \varphi_1 & \varphi_2 & 1 \\ \varphi_1 + \dfrac{\partial \varphi_1}{\partial v}\,dv & \varphi_2 + \dfrac{\partial \varphi_2}{\partial v}\,dv & 1 \\ \varphi_1 + \dfrac{\partial \varphi_1}{\partial u}\,du & \varphi_2 + \dfrac{\partial \varphi_2}{\partial u}\,du & 1 \end{vmatrix} = \pm\frac{1}{2}\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial v} & \dfrac{\partial \varphi_2}{\partial v} \\ \dfrac{\partial \varphi_1}{\partial u} & \dfrac{\partial \varphi_2}{\partial u} \end{vmatrix} du\,dv.$$

Since the area of the curvilinear quadrilateral is nearly equal to twice the area of the triangle, one can write an approximate relation

$$dA = \pm J(u, v)\,du\,dv, \tag{46-3}$$

where

$$J(u, v) \equiv \begin{vmatrix} \dfrac{\partial \varphi_1}{\partial u} & \dfrac{\partial \varphi_2}{\partial u} \\ \dfrac{\partial \varphi_1}{\partial v} & \dfrac{\partial \varphi_2}{\partial v} \end{vmatrix}.$$

The determinant J has many uses in mathematics and is called the *functional determinant or Jacobian.*†

Consider next the double integral

$$\int_R f(x, y)\,dA,$$

---

* The reader, not familiar with the determinantal expression for the area of the triangle, will have no difficulty in deducing it from elementary considerations of analytic geometry

† See Chap. XII.

over some region R, where $f(x, y)$ is continuous in R. This integral may be expressed in terms of the curvilinear coordinates (u, v), by substituting for x and y from (46-2) and by making use of (46-3). Thus,

$$\int_R f(x, y)\, dA = \int_R F(u, v)|J|\, du\, dv, \tag{46-4}$$

where

$$F(u, v) \equiv f[\varphi_1(u, v),\ \varphi_2(u, v)].$$

Some question must be raised regarding the legitimacy of replacing the exact element of area dA by its approximation given by (46-3), but the exact element of area is equal to

$$[|J| + \epsilon]\, du\, dv,$$

where $\epsilon \rightarrow 0$ with du and dv.
The infinitesimal ϵ may be neglected, since in forming the limit of the sum

$$\sum f(x, y)\, \Delta A = \sum F(u, v)|J|\ du\, dv + \epsilon \sum F(u, v)\, du\, dv$$

one can choose du and dv so small that ϵ (and, consequently, the second term in the right-hand member of the foregoing expression) is made as small as desired.*

Setting $f(x, y)$ in (46-4) equal to unity gives the expression for the area of the region R, namely,

$$\int_R dA = \int_R |J|\, du\, dv. \tag{46-5}$$

In evaluating the double integrals in (46-4) and (46-5) by means of iterated integrals, the limits for u and v must be determined from a consideration of the region R, as is shown in the following example:

Example. Let it be required to find the moment of inertia of the area of the circle (Fig. 39)

$$x^2 + y^2 - ax = 0,$$

---

* A more elegant mode of exhibiting the relationship between the elements of area in cartesian and curvilinear coordinates is given in the next chapter. However, the demonstration given above has the advantage of directness of attack on the problem, even though it calls for a justification of the approximation involved

about a diameter of the circle. It is convenient to introduce the polar coordinates

$$x = \rho \cos \theta$$
$$y = \rho \sin \theta,$$

so that the equation of the circle becomes

$$\rho = a \cos \theta.$$

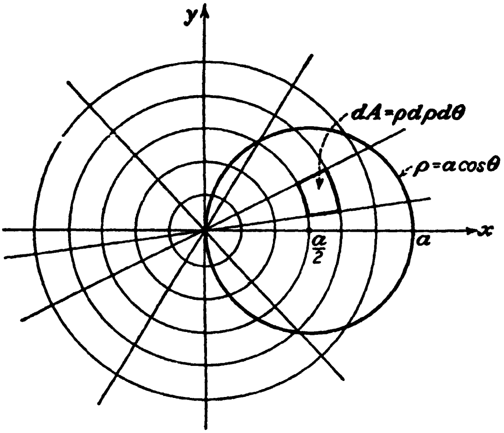

Fig. 39.

Calculating the determinant J gives

$$J = \begin{vmatrix} \cos \theta & \sin \theta \\ -\rho \sin \theta & \rho \cos \theta \end{vmatrix} = \rho,$$

so that

$$dA = \rho \, d\rho \, d\theta.$$

Therefore,

$$\begin{aligned} I = \int_R y^2 \, dA &= \int_0^\pi \int_0^{a \cos \theta} \rho^2 \sin^2 \theta \, \rho \, d\rho \, d\theta \\ &= \int_0^\pi \frac{a^4 \cos^4 \theta \sin^2 \theta}{4} \, d\theta \\ &= \frac{\pi a^4}{64}. \end{aligned}$$

47. Transformation of Points. The discussion in Sec. 46 of the change of variables in a double integral leading to the formula

$$dA = |J|\,du\,dv, \tag{47-1}$$

is capable of a somewhat different interpretation. The expression (47-1) represents the element of area in the curvilinear coordinates defined by the equations

$$\begin{cases} u = f_1(x, y), \\ v = f_2(x, y), \end{cases} \tag{47-2}$$

where x and y are thought to be the ordinary cartesian coordinates. Now the Eqs. (47-2) can be interpreted as the equations of transformation of points from one cartesian coordinate system x, y to another cartesian system u, v.

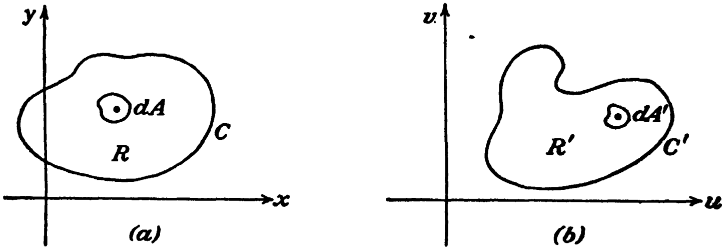

FIG. 40.

Consider some region R in the xy-plane (Fig. 40), and assume that the correspondence of points between the uv-plane and the xy-plane is one to one. This means that (47-2) can be solved uniquely for x and y in terms of u and v to give the inverse transformation

$$\begin{cases} x = \varphi_1(u, v), \\ y = \varphi_2(u, v). \end{cases} \tag{47-3}$$

Corresponding to any point (x, y) of the region R the equations (47-2) determine a pair of numbers (u, v), and conversely, equations (47-3) determine a unique pair of numbers (x, y) for every point (u, v) in some region R' of the uv-plane. The boundary C of the region R will be mapped into some curve C' in the uv-plane which will enclose the region R'. The magnitudes of the areas of the regions R and R' will not be equal in general.

The element of area dA in the xy-plane will be transformed into an element of area dA' in the uv-plane, and the ratio of the magnitudes of these elements of area is precisely equal to the numerical value of the Jacobian. Therefore [see (46-5)],

$$\int_R dy\,dx = \int_{R'} |J|\,du\,dv. \tag{47-4}$$

The integral $\int_{R'} du\,dv$ gives the area bounded by the curve C', and the factor $|J|$ introduced in the integrand of (47-4) takes account of the magnification of the region R produced by the transformation (47-2). This point of view is adopted in the treatment of the change of variables given in Sec. 60.

PROBLEMS

1. Evaluate $\int_R e^{-(x^2+y^2)}\,dy\,dx$, where R is the region bounded by the circle $x^2 + y^2 = a^2$. Use polar coordinates. *Ans.* $\pi(1 - e^{-a^2})$.

2. Find the area outside $\rho = a(1 + \cos\theta)$ and inside $\rho = 3a\cos\theta$.

3. Find the coordinates of the center of gravity of the area between $\rho = 2\sin\theta$ and $\rho = 4\sin\theta$.

4. Calculate the elements of area in the uv-coordinate systems which are related to the cartesian coordinate system xy by means of the following equations of transformation:

(*a*) $x = u + a,\ y = v + b$;
(*b*) $x = au,\ y = bv$;
(*c*) $x = u\cos\alpha - v\sin\alpha,\ y = u\sin\alpha + v\cos\alpha$;

where a, b, and α are constants. Interpret your results geometrically.

5. What are the regions of integration in the uv-coordinate systems of Prob. 4 if the region R in the xy-plane is the interior of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1?$$

6. Discuss the curvilinear coordinate system defined by the relations

$$x = u + v,$$
$$y = u - v;$$

and describe the region in the uv-plane corresponding to the square $x = 1$, $x = 2$, $y = 1$, $y = 2$.

7. Discuss the curvilinear coordinate system defined by the relations

$$u = x^2 - y^2,$$
$$v = 2xy.$$

Sketch the curves $u =$ const. and $v =$ const.

8. Find the area of the cardioid $\rho = a(1 + \cos\theta)$ by integrating first with respect to θ and then with respect to ρ. Check your result by calculating the same area by integrating first with respect to ρ and then with respect to θ.

9. Find by double integration the area bounded by $\rho = 2a \cos 2\theta$ by integrating first with respect to θ.

10. Find the center of gravity of one loop of the curve in Prob. 9.

11. Find the center of gravity of the area which is bounded by one loop of the curve $\rho^2 = 2a^2 \cos 2\theta$ and which is exterior to the circle $\rho = a$.

12. Find the moment of inertia of the area in Prob. 11 about the polar axis.

48. Change of Variables in a Triple Integral. Let the cartesian variables x, y, z be connected with the variables u, v, and w by means of the relations

$$\text{(48-1)} \qquad \begin{cases} u = f_1(x, y, z), \\ v = f_2(x, y, z), \\ w = f_3(x, y, z). \end{cases}$$

Just as in the preceding section, it is assumed that the functions u, v, and w are continuous together with their first partial derivatives in some region R of the xyz-space, and that Eqs. (48-1) can be solved for x, y, and z in terms of u, v, and w.

If u, v, and w are assigned fixed values u_0, v_0 and w_0, then

$$\text{(48-2)} \qquad \begin{cases} u_0 = f_1(x, y, z), \\ v_0 = f_2(x, y, z), \\ w_0 = f_3(x, y, z). \end{cases}$$

The equations (48-2) define three surfaces, two of which intersect in a curve, and the third cuts the curve in a point. Thus, a triplet of values (u_0, v_0, w_0) determines a point in space, and one can regard the totality of numbers (u, v, w) as the curvilinear coordinates. The surfaces $u =$ const., $v =$ const., and $w =$ const. are called the *coordinate surfaces.*

The element of volume dV is enclosed by the three pairs of coordinate surfaces (Fig. 41)

$$u = f_1(x, y, z), \qquad v = f_2(x, y, z), \qquad w = f_3(x, y, z),$$

and

$$u + du = f_1(x, y, z), \qquad v + dv = f_2(x, y, z),$$
$$w + dw = f_3(x, y, z),$$

where du, dv, and dw are assumed to be positive constants.

The calculation leading to the establishment of the formula corresponding to (46-3) is entirely analogous to that of Sec. 46, and all one needs to do is to calculate the volume of a parallelepiped that is approximately equal to the element of volume bounded by the coordinate surfaces

$$\begin{aligned} u &= \text{const.},\\ v &= \text{const.},\\ w &= \text{const.},\\ u + du &= \text{const.},\\ v + dv &= \text{const.},\end{aligned}$$

and

$$w + dw = \text{const.}$$

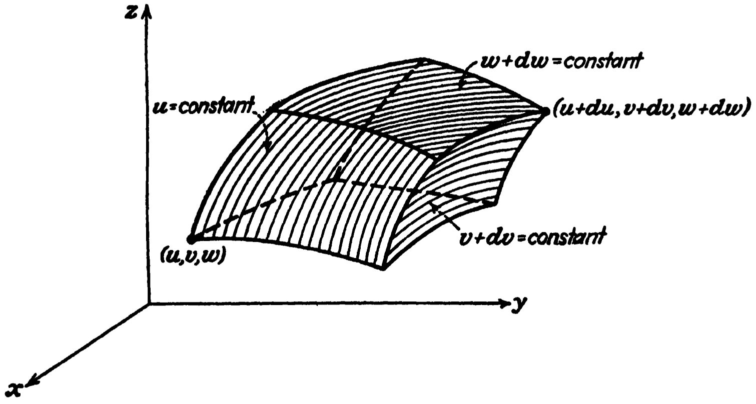

FIG. 41.

If the solutions of (48-1) be denoted by

$$\begin{aligned} x &= \varphi_1(u, v, w),\\ y &= \varphi_2(u, v, w),\\ z &= \varphi_3(u, v, w),\end{aligned}$$

then the element of the volume is

$$dV = |J|\, du\, dv\, dw,$$

where

$$J(u, v, w) \equiv \begin{vmatrix} \dfrac{\partial\varphi_1}{\partial u} & \dfrac{\partial\varphi_2}{\partial u} & \dfrac{\partial\varphi_3}{\partial u} \\ \dfrac{\partial\varphi_1}{\partial v} & \dfrac{\partial\varphi_2}{\partial v} & \dfrac{\partial\varphi_3}{\partial v} \\ \dfrac{\partial\varphi_1}{\partial w} & \dfrac{\partial\varphi_2}{\partial w} & \dfrac{\partial\varphi_3}{\partial w} \end{vmatrix}. \tag{48-3}$$

The Jacobian (48-3) is frequently written in the following forms, which indicate explicitly the variables under consideration,*

$$J\left(\frac{\varphi_1, \varphi_2, \varphi_3}{u, v, w}\right) \quad \text{or} \quad J\left(\frac{x, y, z}{u, v, w}\right).$$

Then

$$\int_R f(x, y, z)\, dV = \int_R F(u, v, w) \cdot \left| J\left(\frac{x, y, z}{u, v, w}\right)\right| \cdot du\, dv\, dw,$$

where

$$F(u, v, w) \equiv f[\varphi_1(u, v, w), \varphi_2(u, v, w), \varphi_3(u, v, w)].$$

An interpretation of the meaning of the Jacobian, analogous to that of Sec. 47, is immediately available if one prefers to think of Eqs. (48-1) as defining the transformation of the points of the region R into the points of the region R' in a cartesian coordinate system (u, v, w). The results of this and the preceding sections can easily be generalized to more than three dimensions.

49. Spherical and Cylindrical Coordinates. Corresponding to the system of polar coordinates in the plane, there are two systems of space coordinates which are frequently used in practical problems. The first of these is the system of spherical, or polar, coordinates. Let $P(x, y, z)$ (Fig. 42) be any point whose projection on the xy-plane is $Q(x, y)$. Then the spherical coordinates of P are ρ, φ, θ, in which ρ is the distance OP, φ is the angle between OQ and the positive x-axis, and θ is the angle

---

* Another notation commonly used to denote the Jacobian is

$$J = \frac{\partial(x, y, z)}{\partial(u, v, w)}.$$

See Chap. XII

between OP and the positive z-axis. Then, from Fig. 42 it is seen that

$$\begin{aligned} x &= OQ\cos\varphi = OP\cos(90^\circ - \theta)\cos\varphi = \rho\sin\theta\cos\varphi, \\ y &= OQ\sin\varphi = \rho\sin\theta\sin\varphi, \\ z &= \rho\cos\theta. \end{aligned}$$

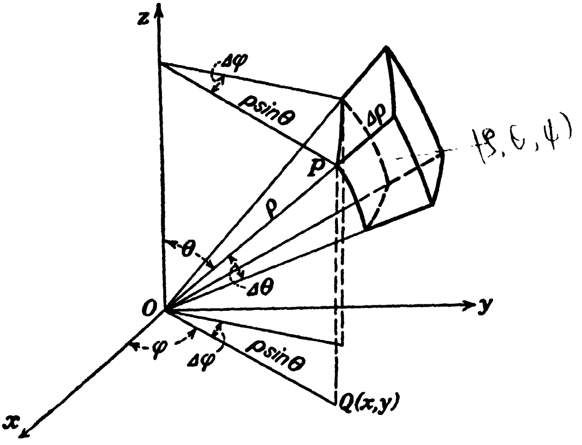

Fig. 42.

The element of volume in spherical coordinates can be obtained by means of (48-3). Since

$$J\left(\frac{x,\ y,\ z}{\rho,\ \varphi,\ \theta}\right) = \begin{vmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ -\rho\sin\theta\sin\varphi & \rho\sin\theta\cos\varphi & 0 \\ \rho\cos\theta\cos\varphi & \rho\cos\theta\sin\varphi & -\rho\sin\theta \end{vmatrix}$$
$$= -\rho^2\sin\theta,$$

it follows that

$$dV = \rho^2 \sin\theta \, d\rho \, d\varphi \, d\theta. \tag{49-1}$$

This element of volume is the volume of the solid bounded by the two concentric spheres of radii ρ and $\rho + d\rho$, the two planes through the z-axis which make angles of φ and $\varphi + d\varphi$ with the xz-plane, and the two cones of revolution whose common axis is the z-axis and whose vertical angles are 2θ and $2(\theta + d\theta)$.

The second space system corresponding to polar coordinates in the plane is the system of cylindrical coordinates. Any point $P(x, y, z)$, whose projection on the xy-plane is Q (Fig. 43), has the cylindrical coordinates ρ, θ, z, where θ is the angle between OQ and the positive x-axis, ρ is the distance OQ, and z is the distance QP. From Fig. 43 it is evident that $x = \rho \cos \theta$, $y = \rho \sin \theta$, and $z = z$. Since

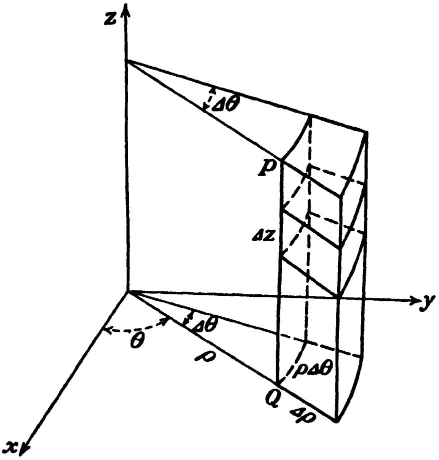

FIG. 43.

$$J\left(\frac{x,\ y,\ z}{\rho,\ \theta,\ z}\right) = \begin{vmatrix} \cos\theta & \sin\theta & 0 \\ -\rho\sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho,$$

it follows that

$$dV = \rho\, d\rho\, d\theta\, dz. \tag{49-2}$$

This element of volume is the volume of the solid bounded by the two cylinders whose radii are ρ and $\rho + d\rho$, the two planes through the z-axis which make angles θ and $\theta + d\theta$ with the xz-plane, and the two planes parallel to the xy-plane at distances z and $z + dz$.

Example 1. Find the x-coordinate of the center of gravity of the solid of uniform density σ lying in the first octant and bounded by the three coordinate planes and the sphere $x^2 + y^2 + z^2 = a^2$.

Since

$$\bar{x} = \frac{\sigma \int_R x\,dV}{\sigma \int_R dV} = \frac{\int_R x\,dV}{\frac{\pi a^3}{6}},$$

it is necessary to compute $\int_R x\,dV$. This integral can be calculated by evaluating the iterated integral

$$\int_0^a \int_0^{\sqrt{a^2-z^2}} \int_0^{\sqrt{a^2-y^2-z^2}} x\,dx\,dy\,dz,$$

but it is easier to transform to spherical coordinates. Then,

$$\begin{aligned}\int_R x\,dV &= \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \int_0^a \rho \sin\theta \cos\varphi \cdot \rho^2 \sin\theta\,d\rho\,d\theta\,d\varphi \\ &= \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{\rho^4}{4}\bigg|_0^a \sin^2\theta \cos\varphi\,d\theta\,d\varphi \\ &= \frac{a^4\pi}{16}\int_0^{\frac{\pi}{2}} \cos\varphi\,d\varphi = \frac{a^4\pi}{16}.\end{aligned}$$

Therefore,

$$\bar{x} = \frac{\frac{a^4\pi}{16}}{\frac{\pi a^3}{6}} = \frac{3a}{8}.$$

Example 2. Let it be required to calculate the moment of inertia I_x of the cylinder used in the illustration of Sec. 45 by transforming the integral into cylindrical coordinates. Then

$$\begin{aligned}I_x = \int_R (y^2 + z^2)\sigma\,dV &= \sigma \int_0^a \int_0^{2\pi} \int_0^b (\rho^2 \sin^2\theta + z^2)\rho\,dz\,d\theta\,d\rho \\ &= \sigma \int_0^a \int_0^{2\pi} \left(\rho^3 z \sin^2\theta + \frac{\rho}{3}z^3\right)\bigg|_0^b d\theta\,d\rho \\ &= \sigma \int_0^a \int_0^{2\pi} \left(b\rho^3 \sin^2\theta + \frac{b^3\rho}{3}\right) d\theta\,d\rho \\ &= \sigma \int_0^a b\left(\frac{\rho^3\theta}{2} - \frac{\rho^3 \sin 2\theta}{4} + \frac{b^2\rho\theta}{3}\right)\bigg|_0^{2\pi} d\rho\end{aligned}$$

$$= \sigma\int_0^a\left(\pi b\rho^3 + \frac{2\pi b^3}{3}\rho\right)d\rho$$

$$= \sigma b\pi\left(\frac{\rho^4}{4} + \frac{b^2\rho^2}{3}\right)\Bigg|_0^a = \frac{\sigma a^2 b\pi}{12}(3a^2 + 4b^2).$$

50. Surface Integrals. Another important application of multiple integrals occurs in defining the area of a surface. The surfaces considered in this section are assumed to have two well-defined sides. If one draws two oppositely directed normals *PN* and *PN'* at any point *P* of a two-sided surface (Fig. 44) and allows the point *P* to move along any path which does not cross the edge of the surface, then the direction of the moving normal *PN* can never be made to coincide with that of *PN'*.

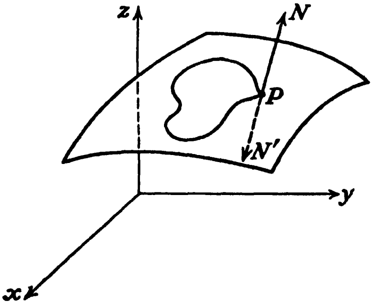

Fig. 44.

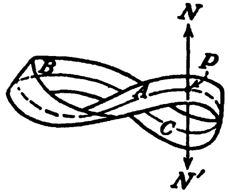

Fig. 45.

There are surfaces that have one side only. Thus a strip of paper glued in such a way that the upper side of one end of the strip is joined onto the under side of the other end will form a one-sided surface (Fig. 45). If two oppositely directed normals *PN* and *PN'* are drawn at any point *P* of this surface, then the normal *PN*, carried along the path *PABCP*, will eventually coincide in direction with *PN'*.

Let $z = f(x, y)$ be the equation of a surface *S* (Fig. 46). The function $z = f(x, y)$ is assumed to have continuous first partial derivatives with respect to x and y. This implies that a continuously turning tangent plane is uniquely defined at every point of the surface *S*.

Let *S'* be a portion of *S* bounded by a closed curve *C*, and such that any line parallel to the z-axis cuts *S'* in only one point. If *C'* is the projection of *C* on the xy-plane, let the region *R*, of which C' is the boundary, be subdivided by lines parallel to the axes into subregions ΔR_i. Through these subdividing lines pass planes parallel to the z-axis. These planes cut from S' small regions $\Delta S'_i$ of area $\Delta\sigma_i$. Let ΔA_i be the area of ΔR_i. Then, except for infinitesimals of higher order,

$$\Delta A_i = \cos\gamma_i\,\Delta\sigma_i,$$

where $\cos\alpha_i$, $\cos\beta_i$, and $\cos\gamma_i$ represent the direction cosines

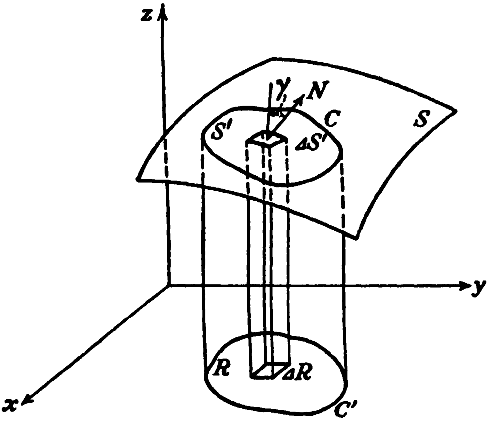

FIG. 46.

of the normal to S at any point (x_i, y_i, z_i) of $\Delta S'_i$. Since (see Sec. 28)

$$\cos\alpha_i:\cos\beta_i:\cos\gamma_i = \left.\frac{\partial z}{\partial x}\right|_i : \left.\frac{\partial z}{\partial y}\right|_i : -1,$$

it follows that

$$\cos\gamma_i = \frac{-1}{\pm\sqrt{\left(\frac{\partial z}{\partial x}\right)_i^2 + \left(\frac{\partial z}{\partial y}\right)_i^2 + 1}}.$$

Using the positive value for $\cos\gamma_i$,

$$\Delta\sigma_i = \sec\gamma_i\,\Delta A_i = \sqrt{\left(\frac{\partial z}{\partial x}\right)_i^2 + \left(\frac{\partial z}{\partial y}\right)_i^2 + 1}\;\Delta A_i.$$

Then

$$\lim_{n\to\infty}\sum_{i=1}^{n}\sqrt{\left(\frac{\partial z}{\partial x}\right)_i^2 + \left(\frac{\partial z}{\partial y}\right)_i^2 + 1}\;\Delta A_i$$

is defined as the area of the surface S'. Since this limit is

$$\int_R \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}\, dA,$$

the value of σ is given by

$$(50\text{-}1)\quad \sigma = \int_R \sec\gamma\, dA = \int\int_R \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}\, dx\, dy.$$

Similarly, by projecting S on the other coordinate planes, it can be shown that

$$\sigma = \int_{R'} \sec\alpha\, dA = \int_{R''} \sec\beta\, dA.$$

The integral of a function $\varphi(x, y, z)$ over the surface $z = f(x, y)$ can now be defined by the equation

$$(50\text{-}2)\quad \int_{S'} \varphi(x, y, z)\, d\sigma = \int\int_R \varphi[x, y, f(x, y)] \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}\, dx\, dy.$$

The outline of the reasoning leading to this definition is as follows: The surface S' is divided into regions $\Delta S'_i$ whose areas are $\Delta\sigma_i$. Let $P_i(x_i, y_i, z_i)$ be any point in $\Delta S'_i$ and form the sum

$$\sum_{i=1}^{n} \varphi(x_i, y_i, z_i)\, \Delta\sigma_i.$$

The limit of this sum as $n \to \infty$, in such a way that every $\Delta\sigma_i \to 0$, is called the surface integral of $\varphi(x, y, z)$ over S' and is denoted by

$$\int_{S'} \varphi(x, y, z)\, d\sigma.$$

This integral can be evaluated with the aid of the formula (50-2).

To ensure the existence of the limit, it is sufficient to assume that the function $\varphi(x, y, z)$ is continuous and single-valued for all points of the surface S'.

Example 1. Find the area of that portion of the surface of the cylinder $x^2 + y^2 = a^2$ which lies in the first octant between the planes $z = 0$ and $z = mx$ (Fig. 47).

This surface can be projected on the xz-plane or on the yz-plane but not on the xy-plane (since any perpendicular to the xy-plane which meets the surface of all will lie on the surface). The projection on the xz-plane is the triangle OAB. Hence

$$\sigma = \int_{OAB} \sec \beta \, dA.$$

But

$$\begin{aligned}\sec \beta &= \sqrt{\left(\frac{\partial y}{\partial x}\right)^2 + 1 + \left(\frac{\partial y}{\partial z}\right)^2} \\ &= \sqrt{\left(\frac{-x}{\sqrt{a^2 - x^2}}\right)^2 + 1 + 0} = a(a^2 - x^2)^{-\frac{1}{2}}.\end{aligned}$$

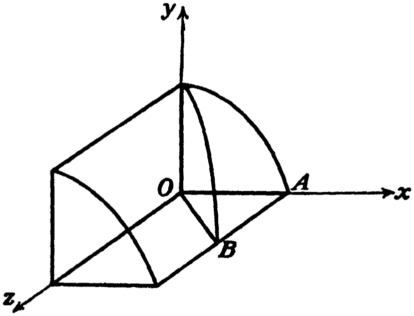

FIG. 47.

Therefore,

$$\begin{aligned}\sigma &= \int_0^a \int_0^{mx} a(a^2 - x^2)^{-\frac{1}{2}} \, dz \, dx \\ &= \int_0^a amx(a^2 - x^2)^{-\frac{1}{2}} \, dx = a^2 m.\end{aligned}$$

Example 2. Find the surface of the sphere $x^2 + y^2 + z^2 = a^2$ cut off by the cylinder $x^2 - ax + y^2 = 0$ (Fig. 48).

From symmetry it is clear that it will suffice to determine the surface in the first octant and multiply the result by 4. Now,

$$\sigma = \int_R \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1} \, dy \, dx,$$

and since $z = \sqrt{a^2 - x^2 - y^2}$,

$$\frac{\partial z}{\partial x} = \frac{-x}{\sqrt{a^2 - x^2 - y^2}}, \qquad \frac{\partial z}{\partial y} = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}.$$

Thus, the integral becomes

$$\begin{aligned}\sigma &= 4\int_0^a\int_0^{\sqrt{ax-x^2}}\sqrt{\frac{x^2+y^2}{a^2-x^2-y^2}+1}\,dy\,dx\\ &= 4\int_0^a\int_0^{\sqrt{ax-x^2}}\frac{a\,dy\,dx}{\sqrt{a^2-x^2-y^2}}.\end{aligned}$$

It is simpler to evaluate this integral by transforming to cylindrical coordinates. The equation of the cylinder becomes $\rho = a\cos\theta$, and that of the sphere

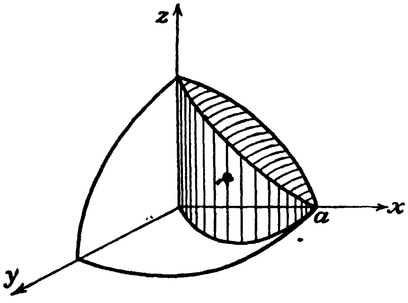

FIG. 48.

$$z = \sqrt{a^2 - x^2 - y^2} = \sqrt{a^2 - \rho^2}.$$

Thus,

$$\sigma = 4\int_0^{\frac{\pi}{2}}\int_0^{a\cos\theta}\frac{a\,\rho\,d\rho\,d\theta}{\sqrt{a^2-\rho^2}} = 4a^2\left(\frac{\pi}{2}-1\right).$$

Example 3. Find the z-coordinate of the center of gravity of one octant of the surface of the sphere $x^2 + y^2 + z^2 = a^2$. Now,

$$\begin{aligned}\bar{z} = \frac{\int_{S'} z\,d\sigma}{\int_{S'} d\sigma} &= \frac{\int_0^a\int_0^{\sqrt{a^2-y^2}} z\sqrt{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2+1}\,dx\,dy}{\frac{4\pi a^2}{8}}\\ &= \frac{2\int_0^a\int_0^{\sqrt{a^2-y^2}} a\,dx\,dy}{\pi a^2} = \frac{a}{2}.\end{aligned}$$

PROBLEMS

1. Find the volume bounded by the cylinder and the sphere of Example 2, page 164. *Ans.* $V = \frac{4}{3}a^3\left(\frac{\pi}{2} - \frac{2}{3}\right)$.

2. Find the surface of the cylinder $x^2 + y^2 = a^2$ cut off by the cylinder $y^2 + z^2 = a^2$. *Ans.* $8a^2$.

3. Find the coordinates of the center of gravity of the portion of the surface of the sphere cut off by the right-circular cone whose vertex is at the center of the sphere. *Ans.* $\bar{x} = a \cos^2 \frac{\alpha}{2}$.

4. Use cylindrical coordinates to find the moment of inertia of the volume of a right-circular cylinder about its axis. *Ans.* $\frac{1}{2}\pi a^4 h$.

5. Find the moments of inertia of the volume of the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

about its axes. *Ans.* $I_z = \frac{4}{15}abc(a^2 + b^2)$.

6. Kinetic energy T is defined as $T = \frac{1}{2}Mv^2$, where M is the mass and v is the velocity of a particle. If the body is rotating with a constant angular velocity ω, show that

$$T = \int_V \tfrac{1}{2}r^2\omega^2 \rho \, dV = \tfrac{1}{2}I\omega^2,$$

where ρ is the density and I is the moment of inertia of the body about the axis of rotation.

7. Find the coordinates of the center of gravity of the area bounded by $x^{1/2} + y^{1/2} = a^{1/2}$, $x = 0$, and $y = 0$.

8. Find the moment of inertia of the area of one loop of $\rho^2 = a^2 \sin 2\theta$ about an axis perpendicular to its plane at the pole.

9. (*a*) Find the expression for dA in terms of u and v, if $x = u(1 - v)$ and $y = uv$;

(*b*) Find the expression for dV in terms of u, v, and w, if $x = u(1 - v)$, $y = uv(1 - w)$, and $z = uvw$.

10. Find the center of gravity of one of the wedges of uniform density cut from the cylinder $x^2 + y^2 = a^2$ by the planes $z = mx$ and $z = -mx$.

11. Find the volume enclosed by the circular cylinder $\rho = 2a \cos \theta$, the cone $z = \rho$, and the plane $z = 0$ (use cylindrical coordinates).

12. Find the center of gravity of the solid of uniform density bounded by the four planes $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, $x = 0$, $y = 0$, and $z = 0$.

13. Find the moment of inertia of the solid of uniform density bounded by the cylinder $x^2 + y^2 = a^2$ and the planes $z = 0$ and $z = b$ about the z-axis.

14. Find, by the method of Sec. 50, the area of the surface of the sphere $x^2 + y^2 + z^2 = a^2$ which lies in the first octant.

15. Prove that

$$J\left(\frac{u,\ v}{x,\ y}\right) \cdot J\left(\frac{x,\ y}{u,\ v}\right) = 1.$$

Hint: Write out the Jacobians and multiply.

16. Prove that

$$J\left(\frac{u,\ v}{x,\ y}\right) \cdot J\left(\frac{x,\ y}{\xi,\ \eta}\right) = J\left(\frac{u,\ v}{\xi,\ \eta}\right),$$

where $u = u(x, y)$, $v = v(x, y)$, $x = x(\xi, \eta)$, and $y = y(\xi, \eta)$.

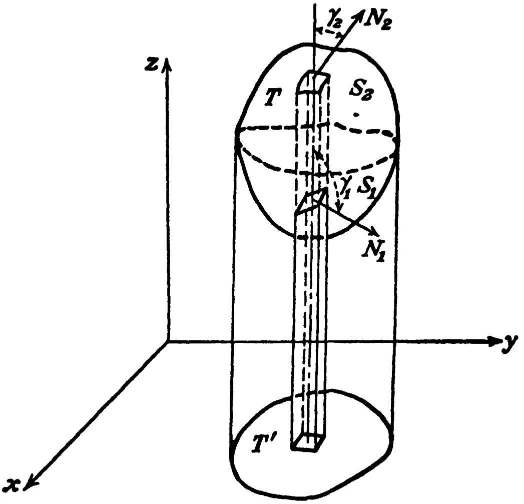

Fig. 49.

51. Green's Theorem in Space.

An important theorem that establishes the connection between the integral over the volume and the integral over the surface enclosing the volume is given next. This theorem has wide applicability in numerous physical problems and is frequently termed the *divergence theorem*.

Theorem. *If $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ and $\frac{\partial P}{\partial x}$, $\frac{\partial Q}{\partial y}$, $\frac{\partial R}{\partial z}$ are continuous and single-valued functions in a region T bounded by a closed surface S, then*

$$\int_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,d\sigma = \int_T \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dV.$$

It will be assumed that S (Fig. 49) is cut by any line parallel to one of the coordinate axes in at most two points. If S is not such a surface, then T is subdivided into regions each of which satisfies this condition, and the extension to more general types of regions presents no difficulty.

A parallel to the z-axis may cut S in two points (x_i, y_i, z_i) and $(x_i, y_i, \bar{z}_i)$, in which $z_i < \bar{z}_i$. Let $z = f_1(x, y)$ be the equation satisfied by (x_i, y_i, z_i) and $z = f_2(x, y)$ be the equation satisfied by $(x_i, y_i, \bar{z}_i)$. Thus S is divided into two parts: S_1, whose equation is $z = f_1(x, y)$; and S_2, whose equation is $z = f_2(x, y)$. Then

$$\int_S R(x, y, z)\cos\gamma\,d\sigma,$$

taken over the exterior of S, is equal to

$$\int_{S_2} R(x, y, z)\cos\gamma\,d\sigma + \int_{S_1} R(x, y, z)\cos\gamma\,d\sigma,$$

taken over the exteriors of the surfaces S_1 and S_2. But, from (50-2), these surface integrals are equal to double integrals taken over the projection T' of T on the xy-plane. Therefore,[*]

$$\begin{aligned}\int_S R(x, y, z)\cos\gamma\,d\sigma &= \int_{T'} \{R[x, y, f_2(x, y)] - R[x, y, f_1(x, y)]\}\,dA\\ &= \int\int_{T'} R(x, y, z)\Big|_{z=f_1(x,y)}^{z=f_2(x,y)} dy\,dx\\ &= \int\int\int_T \frac{\partial R}{\partial z}\,dz\,dy\,dx\end{aligned}$$

or

$$\int_S R(x, y, z)\cos\gamma\,d\sigma = \int_T \frac{\partial R}{\partial z}\,dV.$$

Similarly, it can be shown that

$$\int_S P(x, y, z)\cos\alpha\,d\sigma = \int_T \frac{\partial P}{\partial x}\,dV$$

[*] The negative sign appears in the right-hand member of the equation because

$$\cos\gamma_2\,d\sigma_2 = -\cos\gamma_1\,d\sigma_1,$$

where the subscripts refer to S_2 and S_1.

and

$$\int_S Q(x, y, z) \cos\beta \, d\sigma = \int_T \frac{\partial Q}{\partial y} \, dV.$$

Therefore,

$$\text{(51-1)} \quad \int_S (P \cos\alpha + Q \cos\beta + R \cos\gamma) \, d\sigma = \int_T \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dV.$$

Since $\cos\alpha \, d\sigma = dy \, dz$, $\cos\beta \, d\sigma = dz \, dx$, and $\cos\gamma \, d\sigma = dx \, dy$, (51-1) can be written in the form

$$\text{(51-2)} \quad \int\int_S (P \, dy \, dz + Q \, dz \, dx + R \, dx \, dy) = \int\int\int_T \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx \, dy \, dz.$$

The formula (51-2) bears the name of Green.*

Example. By transforming to a triple integral, evaluate

$$I = \int\int_S (x^3 \, dy \, dz + x^2 y \, dz \, dx + x^2 z \, dx \, dy),$$

where S is the surface bounded by $z = 0$, $z = b$, and $x^2 + y^2 = a^2$.

Calculating the right-hand member with the aid of (51-2) and making use of the symmetry, one finds

$$\begin{aligned} I &= 4\int_0^a \int_0^{\sqrt{a^2-x^2}} \int_0^b (3x^2 + x^2 + x^2) \, dz \, dy \, dx \\ &= 4 \cdot 5b \int_0^a x^2 \sqrt{a^2 - x^2} \, dx \\ &= \tfrac{5}{4} \pi a^4 b. \end{aligned}$$

A direct calculation of the integral I may prove to be instructive. The evaluation of the integral can be carried out by calculating the sum of the integrals evaluated over the projections of the surface S on the coordinate planes. Thus,

$$\begin{aligned} I = &\int_{-a}^a \int_0^b (\sqrt{a^2 - y^2})^3 \, dz \, dy - \int_{-a}^a \int_0^b (-\sqrt{a^2 - y^2})^3 \, dz \, dy \\ &+ \int_{-a}^a \int_0^b x^2 \sqrt{a^2 - x^2} \, dz \, dx - \int_{-a}^a \int_0^b x^2 (-\sqrt{a^2 - x^2}) \, dz \, dx \\ &+ \int_{-a}^a \int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}} (a^2 - y^2) \, b dx \, dy - \int_{-a}^a \int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}} (a^2 - y^2) \cdot 0 \, dx \, dy, \end{aligned}$$

---

* The names of Gauss and Ostrogradsky are also associated with this theorem.

which upon evaluation is seen to check with the result obtained above. It should be noted that the angles α, β, γ are made by the exterior normal with the positive direction of the coordinate axes.

52. Symmetrical Form of Green's Theorem. One of the most widely used formulas in the applications of analysis to a great variety of problems is a form of Green's theorem obtained by setting

$$P = u\frac{\partial v}{\partial x}, \qquad Q = u\frac{\partial v}{\partial y}, \qquad R = u\frac{\partial v}{\partial z}$$

in (51-1). The result of the substitution is

$$\begin{aligned}\int_S u\left(\frac{\partial v}{\partial x}\cos\alpha + \frac{\partial v}{\partial y}\cos\beta + \frac{\partial v}{\partial z}\cos\gamma\right)d\sigma &= \int_T u\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right)dV \\ &\quad + \int_T\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z}\right)dV.\end{aligned}$$

But the direction cosines of the exterior normal n to the surface are

$$\cos\alpha = \frac{dx}{dn}, \qquad \cos\beta = \frac{dy}{dn}, \qquad \cos\gamma = \frac{dz}{dn},$$

so that the foregoing integral reads

$$\text{(52-1)} \quad \begin{aligned}\int_S u\frac{dv}{dn}\,d\sigma &= \int_T u\,\nabla^2 v\,dV \\ &\quad + \int_T\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z}\right)dV,\end{aligned}$$

where

$$\nabla^2 v \equiv \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}.$$

Interchanging the roles of u and v in (52-1) and subtracting the result from (52-1) give the desired formula

$$\int_S\left(u\frac{dv}{dn} - v\frac{du}{dn}\right)d\sigma = \int_T(u\,\nabla^2 v - v\,\nabla^2 u)\,dV.$$

A reference to the conditions imposed upon P, Q, and R in the theorem of Sec. 51 shows that in order to ensure the validity of this formula, it is sufficient to require the continuity of the functions u and v and their first and second space derivatives throughout a closed region T.

PROBLEMS

1. Evaluate, by using Green's theorem,

$$\int\int_S xy\,dy\,dz + y^2\,dz\,dx + yz\,dx\,dy,$$

where S is the surface $x^2 + y^2 + z^2 = a^2$.

2. Show from geometrical considerations that the angle $d\theta$ subtended at the origin by an element ds of a plane curve C is

$$d\theta = \cos(n, r)\frac{ds}{r},$$

where r is the radius vector of the curve, and (n, r) is the angle between the radius vector and the normal to the curve. Hence, show that

$$\theta = \int_C \frac{\cos\,(n, r)\,ds}{r} = \int_C \frac{1}{r}\frac{dr}{dn}\,ds,$$

where the integral is a line integral along the curve C.

3. A solid angle is defined as the angle subtended at the vertex of a cone. The area cut out from a unit sphere by the cone, with its vertex at the center, is called the measure of the solid angle. The measure of the solid angle is clearly equal to the area cut out by the cone from any sphere concentric with the unit sphere divided by the square of the radius of this sphere. In a manner analogous to that employed in Prob. 2 show that the element of solid angle is

$$d\omega = \frac{\cos(n, r)\,d\sigma}{r^2},$$

where the angle between the radius vector and the exterior normal to the surface S is (n, r). Also show that

$$\omega = \int_S \frac{\cos\,(n, r)\,d\sigma}{r^2} = \int_S \frac{1}{r^2}\frac{dr}{dn}\,d\sigma,$$

where the integral is extended over the surface S.

4. By transforming to a triple integral, evaluate

$$\int\int_S x^3\,dy\,dz + y^3\,dz\,dx + z^3\,dx\,dy,$$

where S is the spherical surface $x^2 + y^2 + z^2 = a^2$. Also attempt to calculate this integral directly.

5. Set $v = 1$ in Green's symmetrical formula and assume that u satisfies the equation of Laplace: $\nabla^2 u = 0$. What is the value of $\int_S \frac{du}{dn}\,d\sigma$ if S is an arbitrary closed surface?

6. The density of a square plate varies directly as the square of the distance from one vertex. Find the center of gravity and the moment of inertia of the plate about an axis perpendicular to the plate and passing through the center of gravity.

7. Find the volume of a rectangular hole cut through a sphere if a diameter of the sphere coincides with the axis of the hole.

8. Show that the attraction of a homogeneous sphere at a point exterior to the sphere is the same as though all of the mass of the sphere were concentrated at the center of the sphere. Assume the inverse square law of force.

9. The Newtonian potential V due to a body T at a point P is defined by the equation $V(P) = \int_T \frac{dm}{r}$, where dm is the element of mass of the body and r is the distance from the point P to the element of mass dm. Show that the potential of a homogeneous spherical shell of inner radius b and outer radius a is

$$V = 2\pi\sigma(a^2 - b^2), \qquad \text{if} \qquad r < b,$$

and

$$= \frac{4}{3}\pi\sigma\frac{a^3 - b^3}{r}, \qquad \text{if} \qquad r > a,$$

where σ is the density.

10. Find the Newtonian potential on the axis of a homogeneous circular cylinder of radius a.

11. Show that the force of attraction of a right-circular cone upon a point at its vertex is $2\pi\sigma h(1 - \cos\alpha)$, where h is the altitude of the cone and 2α is the angle at the vertex.

12. Show that the force of attraction of a homogeneous right-circular cylinder upon a point on its axis is

$$2\pi\sigma[h + \sqrt{R^2 + a^2} - \sqrt{(R + h)^2 + a^2}],$$

where h is altitude, a is radius, and R is the distance from the point to one base of the cylinder.

13. Set up the integral representing the part of the surface of the sphere $x^2 + y^2 + z^2 = 100$ intercepted by the planes $x = 1$ and $x = 4$.

14. Find the mass of a sphere whose density varies as the square of the distance from the center.

15. Find the moment of inertia of the sphere in Prob. 14 about a diameter.

16. A circular hole of radius b is drilled through a sphere of radius a. What is the volume drilled out if the axis of the hole coincides with a diameter of the sphere?

17. A circular hole of radius b is drilled through a circular cylinder of radius a in such a way that the axis of the hole coincides with a diameter of the cross section of the cylinder. What is the volume drilled out?

18. Set up the integral for the area of the surface of the sphere intercepted by a right-circular cylinder one of whose elements contains a diameter of the sphere. Evaluate this integral if the radius of the cylinder is equal to one-half the radius of the sphere.

19. If a sphere is inscribed in a right-circular cylinder, then the surfaces of the sphere and the cylinder intercepted by a pair of planes perpendicular to the axis of the cylinder are equal in area. Prove it.

CHAPTER VI

LINE INTEGRALS

53. Definition of Line Integral. Consider a continuous curve C (Fig. 50) joining the points A and B, whose coordinates are (a, b) and (c, d). Let $M(x, y)$ and $N(x, y)$ be two single-valued and continuous* functions of x and y for all points of C. Divide the curve C into n arcs by inserting $n - 1$ points P_i whose coordinates are (x_i, y_i), and define

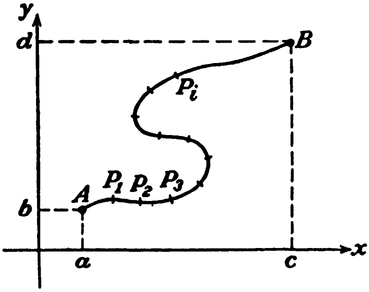

FIG. 50.

$$\Delta x_i \equiv x_i - x_{i-1}$$

and

$$\Delta y_i \equiv y_i - y_{i-1},$$

with $i = 1, 2, \cdots, n$, where $x_0 \equiv a,\ y_0 \equiv b,\ x_n \equiv c,\ y_n \equiv d.$ On each of the n arcs choose a point (ξ_i, η_i), and form the sum

$$\sum_{i=1}^{n} [M(\xi_i, \eta_i)\,\Delta x_i + N(\xi_i, \eta_i)\,\Delta y_i]. \tag{53-1}$$

The limit of this sum as $n \to \infty$ and every Δx_i and Δy_i tends to zero is defined as a line integral along the curve C, or in symbols

$$\lim_{n\to\infty} \sum_{i=1}^{n} [M(\xi_i, \eta_i)\,\Delta x_i + N(\xi_i, \eta_i)\,\Delta y_i] \equiv \int_C [M(x, y)\,dx + N(x, y)\,dy].$$

---

* A function $F(x, y)$ is continuous along the curve C if it is continuous at every point of C. The continuity at a point (x_0, y_0) of the curve C means that $\lim\limits_{\substack{x\to x_0\\ y\to y_0}} F(x, y) = F(x_0, y_0)$ when the point (x, y) approaches (x_0, y_0) along C, that is, $|F(x, y) - F(x_0, y_0)| < \epsilon$ for all points of the curve C that satisfy the inequalities $|x - x_0| < \eta$ and $|y - y_0| < \eta$.

The fact that the limit of the sum exists independently of the mode of subdivision and of the choice of the points (ξ_i, η_i) follows directly from the existence of the ordinary definite integral if the curve C is assumed to be such that it can be broken up into a finite number of monotonic arcs. For, let the equation of the curve C be

$$y = f(x),$$

and assume that C is such that a line parallel to either of the coordinate axes meets the curve in, at most, one point* (Fig. 51).

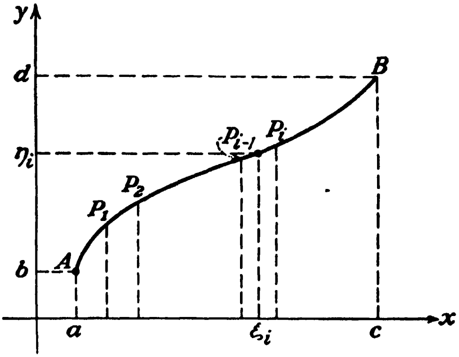

FIG. 51.

Divide this curve into n arcs by inserting the points of subdivision P_i, $(i = 1, 2, \cdots, n - 1)$, the coordinates of which satisfy the conditions

$$a \equiv x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n \equiv c,$$
$$b \equiv y_0 < y_1 < y_2 < \cdots < y_{n-1} < y_n \equiv d,$$

and form the sum

$$\text{(53-2)} \qquad \sum_{i=1}^{n} M(\xi_i, \eta_i)\, \Delta x_i = \sum_{i=1}^{n} M[\xi_i, f(\xi_i)]\, \Delta x_i.$$

Since $M[x, f(x)]$ is a continuous function, the limit of the sum (53-2) as $n \to \infty$ is precisely the ordinary definite integral, so that one can write

$$\lim_{n\to\infty} \sum_{i=1}^{n} M(\xi_i, \eta_i)\, \Delta x_i \equiv \int_a^c M[x, f(x)]\, dx.$$

---

* This is equivalent to saying that the function $f(x)$ is monotone.

Denote the inverse of $y = f(x)$ by $x = \varphi(y)$; then a precisely similar argument shows that

$$\lim_{n\to\infty} \sum_{i=1}^{n} N(\xi_i, \eta_i)\, \Delta y_i = \lim_{n\to\infty} \sum_{i=1}^{n} N[\varphi(\eta_i), \eta_i]\, \Delta y_i$$
$$= \int_b^d N[\varphi(y), y]\, dy.$$

Thus the line integral

$$\int_C [M(x, y)\, dx + N(x, y)\, dy] \tag{53-3}$$

can be expressed as the sum of two definite integrals if the curve C is continuous and monotone. This discussion can be extended in an obvious way if the continuous curve C can be subdivided into a finite number of pieces each of which is monotone.

It is clear from the foregoing that the magnitude of the line integral depends, in general, on the choice of the curve C joining the points A and B, but it should be noted that the definition of the line integral does not require the existence of the derivative of $f(x)$.

If the equation of the curve C is given in a parametric form as

$$x = f_1(t),$$
$$y = f_2(t), \qquad (t_0 \leq t \leq t_1),$$

where $f_1(t)$ and $f_2(t)$ possess continuous derivatives, one can express the line integral (53-3) as a definite integral by substituting for x and y in terms of t. Thus,

$$\int_C [M(x, y)\, dx + N(x, y)\, dy]$$
$$= \int_{t_0}^{t_1} \{M[f_1(t), f_2(t)]f_1'(t) + N[f_1(t), f_2(t)]f_2'(t)\}\, dt,$$

where t_0 and t_1 are the values of the parameter t which correspond to the coordinates of the end points A and B.

If the equation of the curve is given in the form $y = f(x)$, where $f'(x)$ is a continuous function, then $dy = f'(x)\, dx$, and substitution in the line integral (53-3) gives the definite integral

$$\int_C [M(x, y)\, dx + N(x, y)\, dy] = \int_a^c \{M[x, f(x)] + N[x, f(x)]f'(x)\}\, dx.$$

It follows from the properties of definite integrals that the line integral along the curve C from the point A to the point B

is the negative of the line integral from B to A along the same curve.

Example 1. Let the points (0, 0) and (1, 1) be connected by the line $y = x$. Let $M(x, y) = x - y^2$ and $N(x, y) = 2xy$. Then the line integral along $y = x$,

$$I \equiv \int_{(0,0)}^{(1,1)} [(x - y^2)\,dx + 2xy\,dy],$$

becomes, on substitution of $y = x$,

$$\int_0^1 [(x - x^2)\,dx + 2x^2\,dx] = \int_0^1 (x + x^2)\,dx = \tfrac{5}{6}.$$

If (0, 0) and (1, 1) are connected by the parabola $y = x^2$, I along $y = x^2$ is

$$\int_0^1 [(x - x^4)\,dx + 2x^3(2x\,dx)] = \int_0^1 (x + 3x^4)\,dx = \tfrac{11}{10}.$$

Example 2. Consider

$$M(x, y) = 2x^2 + 4xy$$

and

$$N(x, y) = 2x^2 - y^2,$$

with the curve $y = x^2$ connecting the points (1, 1) and (2, 4). Then

$$\int_{(1,1)}^{(2,4)} (M\,dx + N\,dy) = \int_1^2 (2x^2 + 4x \cdot x^2)\,dx + \int_1^4 (2y - y^2)\,dy$$
$$= 13\tfrac{2}{3}.$$

Inasmuch as $dy = 2x\,dx$, this integral can be written as

$$\int_1^2 (2x^2 + 4x^3)\,dx + \int_1^2 (2x^2 - x^4)2x\,dx = 13\tfrac{2}{3}.$$

If the equation of the parabola in this example is written in a parametric form as

$$x = t,$$
$$y = t^2, \qquad (1 \le t \le 2),$$

then the integrand of the line integral can be expressed in terms of the parameter t. Substituting for x, y, dx, and dy in terms of t gives

$$\int_{(1,1)}^{(2,4)} [M\,dx + N\,dy] = \int_1^2 [(2t^2 + 4t^3) + (2t^2 - t^4)2t]\,dt$$
$$= \int_1^2 (2t^2 + 8t^3 - 2t^5)\,dt = 13\tfrac{2}{3}.$$

The reader will readily verify that the value of this integral over a rectilinear path C joining the points (1, 1) and (2, 4) is also $13\frac{4}{5}$. In fact, the value of this integral depends only on the end points and not upon the curve joining them. The reason for this remarkable behavior will appear in Sec. 56.

PROBLEMS

1. Find the value of $\int_{(0,0)}^{(1,1)} [\sqrt{y}\, dx + (x - y)\, dy]$ along the following curves:

(*a*) straight line $x = t$, $y = t$;
(*b*) parabola $x = t^2$, $y = t$;
(*c*) parabola $x = t$, $y = t^2$;
(*d*) cubical parabola $x = t$, $y = t^3$.

2. Find the value of $\int_{(0,0)}^{(1,3)} [x^2y\, dx + (x^2 - y^2)\, dy]$ along (*a*) $y = 3x^2$; (*b*) $y = 3x$.

3. Find the value of $\int_{(0,0)}^{(1,1)} (x^2\, dx + y^2\, dy)$ along the curves of Prob. 1 above.

4. Find the value of $\int_{(0,0)}^{(1,1)} [(x^2 + y^2)\, dx - 2xy\, dy]$ along (*a*) $y = x$; (*b*) $x = y^2$; (*c*) $y = x^2$.

5. Find the value of $\int_{(0,0)}^{(x,y)} (y \sin x\, dx - x \cos y\, dy)$ along $y = x$.

6. Find the value of $\int_{(-a,0)}^{(a,0)} (x\, dy + y\, dx)$ along the upper half of the circle $x^2 + y^2 = a^2$.

7. Evaluate the integral of Prob. 6 over the path formed by the lines $x = -a$, $y = a$, $x = a$. What is the value of this integral if the path is a straight line joining the points $(-a, 0)$ and $(a, 0)$?

8. Find the value of $\int_{(1,0)}^{(0,1)} (x^2\, dx + y^2\, dy)$ along the path given by $x = \sin t$, $y = \cos t$.

9. Evaluate the integral of Prob. 8 if the path is a straight line joining (0, 1) and (1, 0).

10. What is the value of the integral of Prob. 8 if the path is the curve $y = 1 - x^2$?

54. Area of a Closed Curve. Let C be a continuous closed curve which nowhere crosses itself. The equation of such a curve, in parametric form, can be given as

$$\begin{cases} x = f_1(t), \\ y = f_2(t), \end{cases} \tag{54-1}$$

where the parameter t varies continuously from some value $t = t_0$ to $t = t_1$ and the functions $f_1(t)$ and $f_2(t)$ are continuous and single-valued in the interval $t_0 \leq t \leq t_1$. Inasmuch as the curve is assumed to be closed, the initial and the final points of the curve coincide, so that,

$$f_1(t_0) = f_1(t_1),$$

and

$$f_2(t_0) = f_2(t_1).$$

The statement that the curve C does not cut itself implies that there is no other pair of values of the parameter t for which

$$f_1(t') = f_1(t''),$$

and

$$f_2(t') = f_2(t'').$$

A closed curve satisfying the condition stated above will be called *simple.*

As t varies continuously from t_0 to t_1, the points (x, y) determined by (54-1) will trace out the curve C in a certain sense. If C is described so that a man walking along the curve in the direction of the description has the enclosed area always to his left, the curve C is said to be described in the *positive direction,* and the enclosed area will be considered positive; whereas if C is described so that the enclosed area is to the right, then C is described in the *negative direction,* and the area is regarded as negative.

Consider at first a simple closed curve C such that no line parallel to either of the coordinate axes intersects C in more than two points. Let C be bounded by the lines $x = a_1$, $x = a_2$, $y = b_1$, $y = b_2$, which are tangent to C at A_1, A_2, B_1, and B_2, respectively. Clearly C cannot be the graph of a single-valued function. Therefore, let the equation of $A_1B_1A_2$ be given by

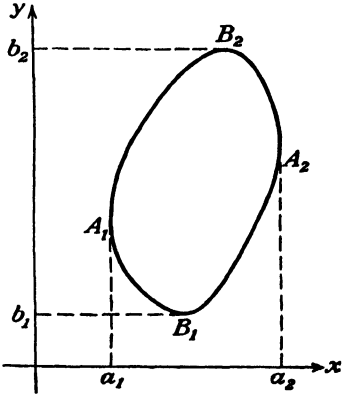

FIG. 52.

$y_1 = f_1(x)$, and the equation of $A_1B_2A_2$ by $y_2 = f_2(x)$, where $f_1(x)$ and $f_2(x)$ are single-valued functions. Then the area enclosed by C (Fig. 52) is given by

$$(54\text{-}2)\qquad \begin{aligned} A &= \int_{a_1}^{a_2} y_2\,dx - \int_{a_1}^{a_2} y_1\,dx \\ &= -\int_{a_2}^{a_1} y_2\,dx - \int_{a_1}^{a_2} y_1\,dx, \end{aligned}$$

or

$$(54\text{-}3)\qquad A = -\int_C y\,dx,$$

in which the last integral is to be taken around C in a counterclockwise direction.

Similarly, if $x_1 = \varphi_1(y)$ is the equation of $B_1A_1B_2$ and $x_2 = \varphi_2(y)$ is the equation of $B_1A_2B_2$:

$$\begin{aligned} A &= \int_{b_1}^{b_2} x_2\,dy - \int_{b_1}^{b_2} x_1\,dy \\ &= \int_{b_1}^{b_2} x_2\,dy + \int_{b_2}^{b_1} x_1\,dy \end{aligned}$$

or

$$(54\text{-}4)\qquad A = \int_C x\,dy.$$

Again, the last integral is to be taken around C in a counterclockwise direction. It may be noted that (54-3) and (54-4) both require that the area be to the left as C is described if the value of A is to be positive.

By adding (54-3) and (54-4), a new formula for A is obtained, namely,

$$(54\text{-}5)\qquad A = \tfrac{1}{2}\int_C (-y\,dx + x\,dy).$$

This formula gives a line-integral expression for A.

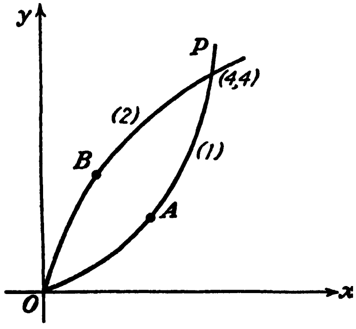

Fig. 53.

To illustrate the application of (54-5), the area between (1) $x^2 = 4y$ and (2) $y^2 = 4x$ (Fig. 53) will be determined.

Then

$$A = \frac{1}{2}\int_C(-y\,dx + x\,dy) = \frac{1}{2}\int_{(1)}(-y\,dx + x\,dy) + \frac{1}{2}\int_{(2)}(-y\,dx + x\,dy)$$

$$= \frac{1}{2}\int_0^4\left(-\frac{x^2}{4}\,dx + x\cdot\frac{x}{2}\,dx\right) + \frac{1}{2}\int_4^0\left(-y\cdot\frac{y}{2}\,dy + \frac{y^2}{4}\,dy\right)$$

$$= \frac{x^3}{24}\bigg|_0^4 - \frac{y^3}{24}\bigg|_4^0 = \frac{16}{3}.$$

For convenience the first integral was expressed in terms of x, whereas the second integral is simpler in terms of y.

The restriction that the curve C be such that no line parallel to either of the coordinate axes cuts it in more than two points can be removed if it is possible to draw a finite number of lines connecting pairs of points on C, so that the area enclosed by the curve is subdivided into regions each of which is of the type considered in the foregoing. This extension is indicated in detail in the following section.

PROBLEMS

1. Find, by using (54-5), the area of the ellipse $x = a\cos\varphi$, $y = b\sin\varphi$.
2. Find, by using (54-5), the area between $y^2 = 9x$ and $y = 3x$.
3. Find, by using (54-5), the area of the hypocycloid of four cusps $x = a\cos^3\theta$, $y = a\sin^3\theta$.
4. Find, by using (54-5), the area of the triangle formed by the line $x + y = a$ and the coordinate axes.
5. Find, by using (54-5), the area enclosed by the loop of the strophoid

$$x = \frac{1 - t^2}{1 + t^2}, \qquad y = \frac{t(1 - t^2)}{1 + t^2}.$$

55. Green's Theorem for the Plane. This remarkable theorem establishes the connection between a line integral and a double integral.

Theorem. *If $M(x, y)$ and $N(x, y)$, $\frac{\partial M}{\partial y}$ and $\frac{\partial N}{\partial x}$ are continuous, single-valued functions over a closed region R, bounded by the curve C, then*

$$\iint_R\left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right)dx\,dy = -\int_C(M\,dx + N\,dy).$$

The double integral is taken over the given region, and the curve C is described in the positive direction.

The theorem will be proved first for a simple closed curve of the type considered in Sec. 54 (see Fig. 52).

Again let $y_1 = f_1(x)$ be the equation of $A_1B_1A_2$ and $y_2 = f_2(x)$ be the equation of $A_1B_2A_2$. Then

$$\begin{aligned}\int\int_R \frac{\partial M}{\partial y}\,dx\,dy &= \int_{a_1}^{a_2} dx \int_{y_1}^{y_2} \frac{\partial M}{\partial y}\,dy \\ &= \int_{a_1}^{a_2} [M(x, y_2) - M(x, y_1)]\,dx \\ &= -\int_{a_2}^{a_1} M(x, y_2)\,dx - \int_{a_1}^{a_2} M(x, y_1)\,dx,\end{aligned}$$

or

$$\int\int_R \frac{\partial M}{\partial y}\,dx\,dy = -\int_C M(x, y)\,dx. \tag{55-1}$$

Similarly, if $x_1 = \varphi_1(y)$ is the equation of $B_1A_1B_2$ and $x_2 = \varphi_2(y)$ is the equation of $B_1A_2B_2$,

$$\begin{aligned}\int\int_R \frac{\partial N}{\partial x}\,dx\,dy &= \int_{b_1}^{b_2} dy \int_{x_1}^{x_2} \frac{\partial N}{\partial x}\,dx = \int_{b_1}^{b_2} [N(x_2, y) - N(x_1, y)]\,dy \\ &= \int_{b_1}^{b_2} N(x_2, y)\,dy + \int_{b_2}^{b_1} N(x_1, y)\,dy,\end{aligned}$$

or

$$\int\int_R \frac{\partial N}{\partial x}\,dx\,dy = \int_C N(x, y)\,dy. \tag{55-2}$$

Therefore, if (55-2) is subtracted from (55-1),

$$\int\int_R \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) dx\,dy = -\int_C [M(x, y)\,dx + N(x, y)\,dy]. \tag{55-3}$$

It will be observed that setting $M = -y$ and $N = x$ gives the formula (54-5).

Now let the region have any continuous boundary curve C, so long as it is possible to draw a finite number of lines which divide the region into subregions each of the type considered in the first part of this section; that is, the subregions must have boundary curves which are cut by any parallel to either of the coordinate axes in at most two points. Such a region R is shown in Fig. 54.

By drawing the lines A_1A_2 and A_3A_4, the region R is divided into three subregions R_1, R_2, and R_3. The boundary curve of each region is of the simple type. The positive direction of each boundary curve is indicated by the arrows. The theorem can be applied to each subregion separately. When the three equations are added, the left-hand members add to give the double integral over the entire region R. The right-hand members give

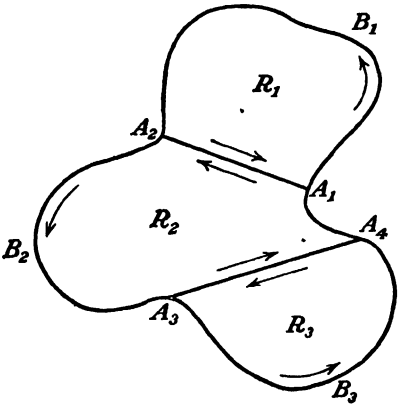

Fig. 54.

$$-\int_{C_1}[M\,dx + N\,dy] - \int_{C_2}[M\,dx + N\,dy] - \int_{C_3}[M\,dx + N\,dy],$$

where

$$\begin{aligned} C_1 &= A_1B_1A_2 + A_2A_1, \\ C_2 &= A_1A_2 + A_2B_2A_3 + A_3A_4 + A_4A_1, \\ C_3 &= A_4A_3 + A_3B_3A_4. \end{aligned}$$

Since each of the lines A_1A_2 and A_3A_4 is traversed once in each direction, the line integrals which arise from them will cancel. The remaining line integrals, taken over the arcs of C, add to give the line integral over C. Therefore,

$$\int\int_R\left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) dx\,dy = -\int_C(M\,dx + N\,dy)$$

holds for regions of the type R.

Another type of region in which an auxiliary line is introduced is the region whose boundary is formed by two or more distinct curves. Thus, if R (Fig. 55) is the region between C_1 and C_2, the line A_1A_2 is drawn in order to make the total boundary

$$C_1 + A_2A_1 + C_2 + A_1A_2$$

a single curve. The theorem can be applied, and the line integrals over A_2A_1 and A_1A_2 will cancel, leaving only the line integrals over C_1 and C_2.

If the region R is such that any closed curve drawn in it can, by a continuous deformation, be shrunk to a point without crossing the boundary of the region, then the latter is called *simply connected*. Thus, regions bounded by a circle, a rectangle, or an ellipse are simply connected. The region R exterior to C_2 and interior to C_1 (Fig. 55) is not simply connected because a circle drawn within R and enclosing C_2 cannot be shrunk to a point without crossing C_2. In ordinary parlance, regions that have holes are not simply connected regions; they are called *multiply connected regions*. The importance of this classification will appear in the next two sections.

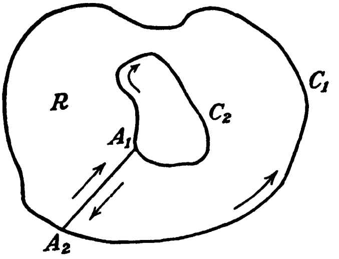

FIG. 55.

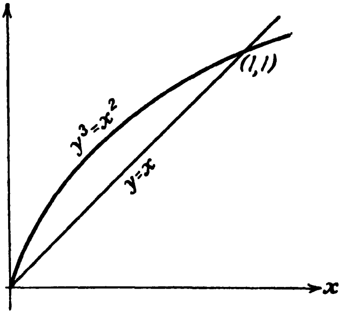

FIG. 56.

Example. Evaluate by using Green's theorem

$$\int_C (x^2y\,dx + y^3\,dy),$$

where C is the closed path formed by $y = x$ and $y^3 = x^2$ from (0, 0) to (1, 1) (Fig. 56). Since $M = x^2y$ and $N = y^3$,

$$\frac{\partial M}{\partial y} = x^2 \qquad \text{and} \qquad \frac{\partial N}{\partial x} = 0.$$

Then,

$$\begin{aligned}\int_C (x^2y\,dx + y^3\,dy) &= -\int\int_R \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) dx\,dy \\ &= -\int_0^1 \int_{y^{3/2}}^{y} x^2\,dx\,dy = -\int_0^1 \frac{x^3}{3}\bigg|_{y^{3/2}}^{y} dy \\ &= \int_0^1 \left(\frac{y^{9/2}}{3} - \frac{y^3}{3}\right) dy = -\frac{1}{44}.\end{aligned}$$

PROBLEMS

1. Find, by Green's theorem, the value of

$$\int_C (x^2y\,dx + y\,dy)$$

along the closed curve C formed by $y^2 = x$ and $y = x$ between (0, 0) and (1, 1).

2. Find, by Green's theorem, the value of

$$\int_C [(x^2 + y)\,dx + (x - y^2)\,dy]$$

along the closed curve C formed by $y^3 = x^2$ and $y = x$ between (0, 0) and (1, 1).

3. Use Green's theorem to find the value of

$$\int_C [(xy - x^2)\,dx + x^2y\,dy]$$

along the closed curve C formed by $y = 0$, $x = 1$, and $y = x$.

4. Use Green's theorem to evaluate

$$\int_C \left(\frac{1}{y}\,dx + \frac{1}{x}\,dy\right)$$

along the closed path formed by $y = 1$, $x = 4$, and $y = +\sqrt{x}$.

5. Check the answers of the four preceding problems by evaluating the line integrals directly.

56. Properties of Line Integrals.

Theorem 1. *Let M and N be two functions of x and y, such that M, N, $\frac{\partial M}{\partial y}$ and $\frac{\partial N}{\partial x}$ are continuous and single-valued at every point of a simply connected region R. The necessary and sufficient* *condition that* $\int_C (M\,dx + N\,dy) = 0$ *around every closed curve* C *drawn in* R, *is that*

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x},$$

for every point of R.

Since

$$-\int_C (M\,dx + N\,dy) = \int\int_A \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) dx\,dy,$$

where A is the region enclosed by C, it follows that

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

makes the double integral, and consequently the line integral, have the value zero. Conversely, let $\int_C (M\,dx + N\,dy) = 0$ around every closed curve C drawn in R. Suppose that

$$\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} \neq 0$$

at some point P of R. Since $\frac{\partial M}{\partial y}$ and $\frac{\partial N}{\partial x}$ are continuous functions of x and y,

$$\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}$$

is also a continuous function of x and y. Therefore, there must exist some region S about P, in which $\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}$ has the same sign as at P. Then

$$\int\int_S \left(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}\right) dx\,dy \neq 0,$$

and hence $\int (M\,dx + N\,dy) \neq 0$ around the boundary of this region. This contradicts the hypothesis that

$$\int_C (M\,dx + N\,dy) = 0$$

around every closed curve C drawn in R. It follows that

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

at all points of R.

Example 1. Let

$$M = \frac{-y}{x^2 + y^2} \qquad \text{and} \qquad N = \frac{x}{x^2 + y^2}.$$

Then

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

M, N, $\dfrac{\partial M}{\partial y}$ and $\dfrac{\partial N}{\partial x}$ are continuous and single-valued for all points of the xy-plane except (0, 0). Hence $\int_C (M\,dx + N\,dy) = 0$ around any closed curve C (Fig. 57) which does not enclose (0, 0). In polar coordinates, obtained by the change of variables,

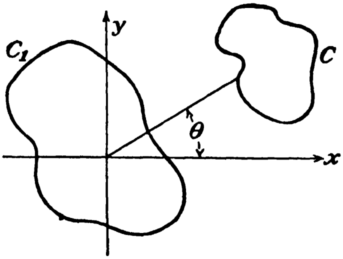

FIG. 57.

$$x = \rho \cos \theta, \qquad y = \rho \sin \theta,$$

$$\int_C \left(\frac{-y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy \right) = \int_C d\theta.$$

If C does not enclose the origin, θ varies along C from its original value θ_0 back to θ_0. Therefore, $\int_C d\theta = 0$. If C_1 encloses the origin, θ varies along C_1 from θ_0 to $\theta_0 + 2\pi$, so that $\int_{C_1} d\theta = 2\pi$.

Example 2. Find, by Green's theorem, the value of

$$I = \int_C [(x^2 + xy)\,dx + (y^2 + x^2)\,dy],$$

where C is the square formed by the lines $y = \pm 1$ and $x = \pm 1$. Since

$$\frac{\partial M}{\partial y} = x, \qquad \frac{\partial N}{\partial x} = 2x,$$

$$I = \int_{-1}^{1} \int_{-1}^{1} x\,dy\,dx = 0$$

Note that the line integral has the value zero, but $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$. This does not contradict Theorem 1. Why?

Theorem 2. *Let M and N satisfy the conditions of Theorem* 1. *The necessary and sufficient condition that* $\int_{(a,b)}^{(x,y)} (M\,dx + N\,dy)$ *be independent of the curve connecting* (a, b) *and* (x, y) *is that* $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ *at all points of the region R. In this case the line integral is a function of the end points only.*

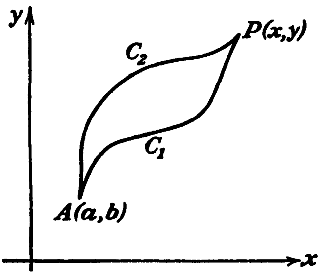

FIG. 58.

Suppose $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Let C_1 and C_2 (Fig. 58) be any two curves from A to P, and let

$$I_1 \equiv \int_{C_1} (M\,dx + N\,dy)$$

and

$$I_2 \equiv \int_{C_2} (M\,dx + N\,dy)$$

be the values of the line integral from A to P along C_1 and C_2, respectively. Then $I_1 - I_2$ is the value of the integral around the closed path formed by C_1 and C_2. By Theorem 1,

$$I_1 - I_2 = 0.$$

Therefore, $I_1 = I_2$, so that the line integral taken over any two paths from A to P has the same value.

Conversely, suppose that $\int (M\,dx + N\,dy)$ is independent of the path from A to P. Then, for any two curves C_1 and C_2, $I_1 = I_2$. It follows that $\int (M\,dx + N\,dy) = 0$ for the closed path formed by C_1 and C_2. Hence, by Theorem 1, $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$.

Example. Consider

$$\int_{(1,1)}^{(2,2)} \left(\frac{1 + y^2}{x^3}\,dx - \frac{1 + x^2}{x^2} y\,dy \right).$$

Since $\frac{\partial M}{\partial y} = \frac{2y}{x^3}$ and $\frac{\partial N}{\partial x} = \frac{2y}{x^3}$, and both functions are continuous except at $(0, 0)$, the line integral is independent of the path so long as it does not enclose the origin. Choose $y = 1$ from $(1, 1)$ to $(2, 1)$, and $x = 2$ from $(2, 1)$ to $(2, 2)$, as the path of integration. Then

$$I = \int_1^2 \frac{2}{x^3}\,dx - \int_1^2 \frac{5}{4}y\,dy = -\frac{1}{x^2}\bigg|_1^2 - \frac{5}{8}y^2\bigg|_1^2 = -\frac{9}{8}.$$

Theorem 3. *Let M and N satisfy the conditions of Theorem 1. The necessary and sufficient condition that there exist a function $F(x, y)$ such that $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$ is that $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ at all points of the region R.*

If $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, Theorem 2 proves that

$$\int_{(a,b)}^{(x,y)} (M\,dx + N\,dy)$$

is independent of the path. Therefore,

$$\int_{(a,b)}^{(x,y)} (M\,dx + N\,dy) = F(x, y), \tag{56-1}$$

and this function $F(x, y)$ depends only on the coordinates of the end points of the path. Hence,

$$F(x + \Delta x, y) = \int_{(a,b)}^{(x+\Delta x,y)} (M\,dx + N\,dy).$$

Let the path of integration be chosen as a curve C (Fig. 59) from A to P and the straight line PP' from $P(x, y)$ to $P'(x + \Delta x, y)$. Then,

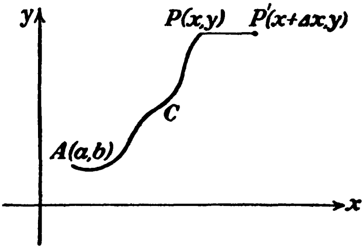

FIG. 59.

$$F(x + \Delta x, y) = \int_{(a,b)}^{(x,y)} (M\,dx + N\,dy) + \int_{(x,y)}^{(x+\Delta x,y)} (M\,dx + N\,dy)$$

or

$$F(x + \Delta x, y) = F(x, y) + \int_x^{x+\Delta x} M(x, y)\,dx. \tag{56-2}$$

The second integral reduces to the simpler form given in (56-2) since y is constant along PP', and therefore $dy = 0$. From (56-2)

$$\frac{\partial F}{\partial x} \equiv \lim_{\Delta x \to 0}\left[\frac{F(x + \Delta x, y) - F(x, y)}{\Delta x}\right]$$
$$= \lim_{\Delta x \to 0}\left[\frac{1}{\Delta x}\int_x^{x+\Delta x} M(x, y)\, dx\right]$$

Application of the first mean-value theorem* gives

$$\int_x^{x+\Delta x} M(x, y)\, dx = \Delta x\, M(\xi, y), \qquad (x \leq \xi \leq x + \Delta x).$$

Therefore,

$$\frac{\partial F}{\partial x} = \lim_{\Delta x \to 0}\left[\frac{1}{\Delta x} \cdot \Delta x\, M(\xi, y)\right] = \lim_{\Delta x \to 0} M(\xi, y).$$

Hence,

$$\frac{\partial F}{\partial x} = M(x, y).$$

It can be proved similarly that

$$\frac{\partial F}{\partial y} = N(x, y).$$

The function F is really a function of both end points. Multiplying $\frac{\partial F}{\partial x} = M(x, y)$ by dx and $\frac{\partial F}{\partial y} = N(x, y)$ by dy gives

$$dF = \frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy = M(x, y)\, dx + N(x, y)\, dy.$$

Thus, if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x},$$

the integrand in $\int_C (M\, dx + N\, dy)$ is the exact differential of the function $F(x, y)$, which is determined by the formula (56-1).

The most general expression for a function $\Phi(x, y)$, whose total differential is $d\Phi = M\, dx + N\, dy$, is

$$\Phi(x, y) = F(x, y) + C,$$

---

[*] See Sec. 37.

where C is an arbitrary constant. Indeed, since dF and $d\Phi$ are equal,

$$d(F - \Phi) = 0,$$

so that

$$F - \Phi = \text{const.}$$

To prove the necessity of the condition of the theorem, note that if there exists a function $F(x, y)$ such that

$$\frac{\partial F}{\partial x} = M(x, y) \qquad \text{and} \qquad \frac{\partial F}{\partial y} = N(x, y),$$

then

$$\frac{\partial^2 F}{\partial y\, \partial x} = \frac{\partial M}{\partial y} \qquad \text{and} \qquad \frac{\partial^2 F}{\partial x\, \partial y} = \frac{\partial N}{\partial x}.$$

Since $\dfrac{\partial M}{\partial y}$ and $\dfrac{\partial N}{\partial x}$ are both continuous, $\dfrac{\partial^2 F}{\partial x\, \partial y}$ and $\dfrac{\partial^2 F}{\partial y\, \partial x}$ are also continuous, and hence,*

$$\frac{\partial^2 F}{\partial x\, \partial y} = \frac{\partial^2 F}{\partial y\, \partial x}.$$

Therefore,

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

As a corollary to Theorem 3, one can state the following: *The necessary and sufficient condition that $M(x, y)\, dx + N(x, y)\, dy$ be an exact differential is that* $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$.

PROBLEMS

1. Show that

$$\int_{(0,1)}^{(1,2)} [(x^2 + y^2)\, dx + 2xy\, dy]$$

is independent of the path, and determine its value.

2. Test for independence of path:

(*a*) $\int (y \cos x\, dx + \sin x\, dy)$;

(*b*) $\int [(x^2 - y^2)\, dx + 2xy\, dy]$;

---

* See Sec. 31.

(*c*) $\int [(x - y^2)\,dx + 2xy\,dy]$;

(*d*) $\int [(x^3 - y^2)\,dx - 2(x - 1)y\,dy]$.

3. Show that $\int_{0,0)}^{(1,1)} \left[\frac{(1 - y^2)}{(1 + x)^3}\,dx + \frac{y}{(1 + x)^2}\,dy \right]$ is independent of the path and find its value.

4. Show that the line integral

$$\int_C \left[\frac{-y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2} \right]$$

evaluated along a square 2 units on the side and with center at the origin has the value 2π. Give the reason for failure of this integral to vanish along this closed path.

5. Find the values of the following line integrals:

(*a*) $\int_{(0,0)}^{\left(\frac{\pi}{2},\frac{\pi}{2}\right)} (y \cos x\,dx + \sin x\,dy)$.

(*b*) $\int_{(0,0)}^{\left(\frac{1}{2},\frac{1}{2}\right)} \left[\frac{xy\,dx}{\sqrt{1 - x^2}} - \sqrt{1 - x^2}\,dy \right]$.

(*c*) $\int_{(1,1)}^{(2,3)} [(x + 1)\,dx + (y + 1)\,dy]$.

57. Multiply Connected Regions. It was shown that the necessary and sufficient condition for the vanishing of the line integral $\int_C [M(x, y)\,dx + N(x, y)\,dy]$ around the closed path C is the equality of $\frac{\partial M}{\partial y}$ and $\frac{\partial N}{\partial x}$ at every point of the region enclosed by C. It was assumed that C was drawn in a simply connected region R and that the functions $M(x, y)$ and $N(x, y)$, together with their first partial derivatives, were continuous on and in the interior of C. The latter condition was imposed in order to ensure the integrability of the functions involved. The reason for imposing the restriction on the connectivity of the region essentially lies in the type of regions permitted by Green's theorem.

Thus consider a region R containing one hole (Fig. 60). The region R will be assumed to consist of the exterior of C_2 and the interior of C . Let a closed contour C be drawn, which lies entirely in R and encloses C_2. Now, even though the functions $M(x, y)$ and $N(x, y)$ together with their derivatives may be continuous in R, the integral $\int_C [M(x, y)\,dx + N(x, y)\,dy]$ may not vanish. For, let K be any other closed curve lying in R and enclosing C_2, and suppose that the points A and B of K and C be joined by a straight line AB Consider the integrals

$$\int_{APA} + \int_{AB} + \int_{BQB} + \int_{BA},$$

where the subscripts on the integrals indicate the direction of integration along the curves K, C, and along the straight line AB, as is indicated in Fig. 60. Since the path AB is traversed twice in opposite directions, the second and the last of the integrals above will annul each other, so that there will remain only the integral along K, traversed in the counterclockwise direction, and the integral along C, in the clockwise direction. Now if M and N satisfy the conditions of Theorem 1, Sec. 56, then

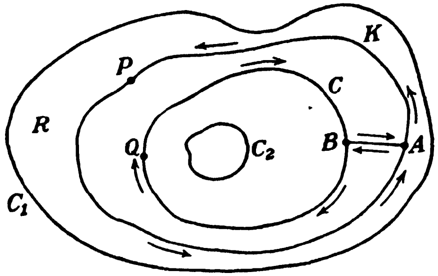

Fig. 60.

$$\int_{\circlearrowleft K} [M\,dx + N\,dy] + \int_{\circlearrowright C} [M\,dx + N\,dy] = 0,$$

where the arrows on the circles indicate the direction of integration. Thus

$$\int_{\circlearrowleft K} [M\,dx + N\,dy] = \int_{\circlearrowleft C} [M\,dx + N\,dy], \tag{57-1}$$

both integrals being taken in the counterclockwise direction.

The important statement embodied in (57-1) is that the magnitude of the line integral evaluated over a closed path in R, surrounding the hole, has the same constant value whatever be the path enclosing C_2. This value need not be zero, as is seen from a simple example already mentioned in Sec. 56. Thus, let the region R consist of the exterior of the circle of radius unity and with center at the origin and of the interior of a concentric circle of radius 3 (Fig. 61). The functions $M = \frac{-y}{x^2 + y^2}$ and $N = \frac{x}{x^2 + y^2}$, and their derivatives, obviously satisfy the conditions of continuity in R and on C_1 and C_2. Also, $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. But

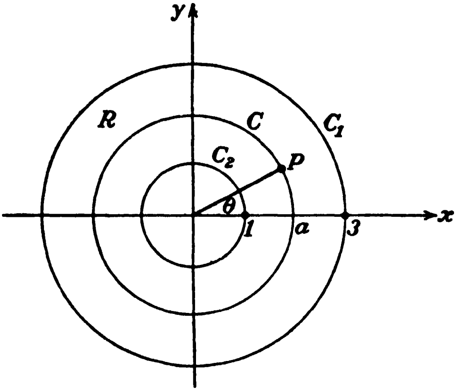

FIG. 61.

$$\int_C \left(\frac{-y\,dx}{x^2 + y^2} + \frac{x}{x^2 + y^2}\,dy \right),$$

where C is the circle

$$x = a\cos\theta,$$
$$y = a\sin\theta, \qquad (1 < a < 3),$$

gives

$$\int_0^{2\pi} \frac{a^2 \sin^2\theta + a^2\cos^2\theta}{a^2}\,d\theta = 2\pi.$$

The function $F(x, y)$, of which $M(x, y)\,dx + N(x, y)\,dy$ is an exact differential, is $F(x, y) = \tan^{-1}\frac{y}{x}$, which is a multiple-valued function.

The function

$$F(x, y) = \int_{(a,b)}^{(x,y)} [M(x, y)\, dx + N(x, y)\, dy],$$

where M and N satisfy the conditions of Theorem 1, Sec. 56, will be single-valued if the region R is simply connected (as is required in Theorem 1) but not necessarily so if the region is multiply connected.

58. Line Integrals in Space. The line integral over a space curve C is defined in a way entirely analogous to that described in Sec. 56.

Let C be a continuous space curve joining the points A and B, and let $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$ be three continuous, single-valued functions of the variables x, y, z. Divide the curve C into n arcs Δs_i $(i = 1, 2, \cdots, n)$, whose projections on the coordinate axes are Δx_i, Δy_i, Δz_i, and form the sum

$$\sum_{i=1}^{n} [P(\xi_i, \eta_i, \zeta_i)\, \Delta x_i + Q(\xi_i, \eta_i, \zeta_i)\, \Delta y_i + R(\xi_i, \eta_i, \zeta_i)\, \Delta z_i],$$

where (ξ_i, η_i, ζ_i) is a point chosen at random on the arc Δs_i. The limit of this sum as n increases indefinitely in such a way that each $\Delta s_i \to 0$ is called the *line integral* of $P\, dx + Q\, dy + R\, dz$, taken along C between the points A and B. It is denoted by the symbol

$$\int_C [P(x, y, z)\, dx + Q(x, y, z)\, dy + R(x, y, z)\, dz]. \tag{58-1}$$

The conditions imposed upon the functions P, Q, and R are sufficient to ensure the existence of the limit, provided that the curve C is suitably restricted.

If the equation of the space curve C is given in parametric form as

$$\begin{cases} x = f_1(t), \\ y = f_2(t), \\ z = f_3(t), \qquad (t_0 \le t \le t_1), \end{cases} \tag{58-2}$$

where $f_1(t)$, $f_2(t)$, and $f_3(t)$ possess continuous derivatives in the interval $t_0 \le t \le t_1$, the line integral (58-1) can be expressed as a definite integral

$$\int_{t_0}^{t_1} [P \cdot f_1'(t) + Q \cdot f_2'(t) + R \cdot f_3'(t)]\, dt$$

where P, Q, and R are expressed in terms of t with the aid of (58-2).

59. Stokes's Theorem. Simply connected regions only will be considered in the remainder of this chapter. A simply connected space region is characterized by the fact that any closed curve drawn in it can be shrunk to a point without crossing the surfaces bounding the region. Thus, a cube, a sphere, a cylinder, or a region between two concentric spheres are examples of simply connected space regions. An anchor ring, or torus, is not simply connected.

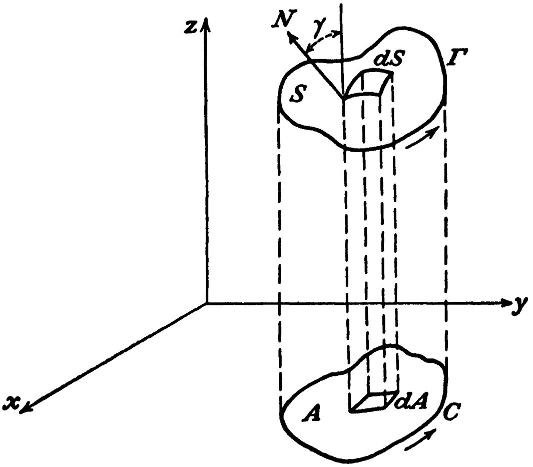

FIG. 62.

In considering closed surfaces, the direction of the exterior normal to the surface will be reckoned as positive. If the surface is open and two-sided (such as that of a hemisphere, for example), the direction of the positive normal may be taken at will. Let an open two-sided surface S be bounded by a closed curve Γ (Fig. 62); then the motion along Γ will be regarded as positive if a man walking along Γ with his head in the direction of the positive normal has the surface S to his left.

Stokes's Theorem. *Let $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$ and their partial derivatives with respect to x, y, and z be continuous and single-valued functions in a region containing the surface S*

which is bounded by the closed curve Γ. *Let* dS *be the element of area of* S, *and let* $\cos\alpha$, $\cos\beta$, *and* $\cos\gamma$ *be the direction cosines of the exterior normal to* dS. *Then*

$$\int_\Gamma (P\,dx + Q\,dy + R\,dz) = \int\int_S \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\cos\beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\cos\gamma\right] dS.$$

Let $z = f(x, y)$ be the equation of S and consider

$$P(x, y, z) = P[x, y, f(x, y)] \equiv M(x, y).$$

Let C be the projection of Γ on the xy-plane, so that

$$\int_\Gamma P(x, y, z)\,dx = \int_C M(x, y)\,dx.$$

From Green's theorem for the plane,

$$\int_C M(x, y)\,dx = -\int\int_A \frac{\partial M}{\partial y}\,dx\,dy,$$

where A is the projection of S on the xy-plane. But

$$\frac{\partial M}{\partial y} = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y},$$

so that

$$\int_\Gamma P(x, y, z)\,dx = -\int\int_A \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y}\right) dx\,dy.$$

Since[*]

$$\frac{\partial z}{\partial x}:\frac{\partial z}{\partial y}: -1 = \cos\alpha : \cos\beta : \cos\gamma, \qquad \frac{\partial z}{\partial y} = -\frac{\cos\beta}{\cos\gamma}.$$

Moreover,

$$dx\,dy = \cos\gamma\,dS.$$

Therefore,

$$\text{(59-1)} \qquad \int_\Gamma P(x, y, z)\,dx = \int\int_S \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right) dS.$$

[*] See Sec. 28.

Similarly, it can be shown that

$$\text{(59-2)} \qquad \int_\Gamma Q\,dy = \int\!\!\int \left(\frac{\partial Q}{\partial x}\cos\gamma - \frac{\partial Q}{\partial z}\cos\alpha\right) dS,$$

$$\text{(59-3)} \qquad \int_\Gamma R\,dz = \int\!\!\int_S \left(\frac{\partial R}{\partial y}\cos\alpha - \frac{\partial R}{\partial x}\cos\beta\right) dS.$$

By adding (59-1), (59-2), and (59-3),

$$\text{(59-4)} \qquad \int_\Gamma (P\,dx + Q\,dy + R\,dz) = \int\!\!\int_S \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\cos\beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\cos\gamma\right] dS.$$

Using the fact that $\cos\alpha\,dS = dy\,dz$, $\cos\beta\,dS = dz\,dx$, and $\cos\gamma\,dS = dx\,dy$, (59-4) can be written in the form

$$\text{(59-5)} \qquad \int_\Gamma (P\,dx + Q\,dy + R\,dz) = \int\!\!\int_S \left[\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy\,dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz\,dx\right].$$

By means of this theorem, it is possible to derive three theorems analogous to those given in Sec. 56 for line integrals in the plane. Since the proofs of these theorems are similar to those given in Sec. 56, they will be omitted here.

Theorem 1. *Let the region of space considered be one in which* $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$ *and their partial derivatives are continuous and single-valued functions of* x, y, *and* z. *Then the necessary and sufficient condition that*

$$\int (P\,dx + Q\,dy + R\,dz) = 0$$

around every closed curve in the region is that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z},$$

for every point of the region.

Theorem 2. *Let the functions considered satisfy the conditions of Theorem* 1. *Then the necessary and sufficient condition that*

$$\int_{(a,b,c)}^{(x,y,z)} (P\,dx + Q\,dy + R\,dz)$$

be independent of the path from (a, b, c) *to* (x, y, z) *is that*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z},$$

for every point of the region.

Theorem 3. *Let the functions* P, Q, *and* R *satisfy the conditions of Theorem* 1. *Then the necessary and sufficient condition that there exist a function* $F(x, y, z)$, *such that*

$$\frac{\partial F}{\partial x} = P, \qquad \frac{\partial F}{\partial y} = Q, \qquad \frac{\partial F}{\partial z} = R,$$

is that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z},$$

for every point of the region. The function $F(x, y, z)$ *is given by the formula*

$$F(x, y, z) = \int_{(a,b,c)}^{(x,y,z)} (P\,dx + Q\,dy + R\,dz).$$

Corollary. *The necessary and sufficient condition that*

$$P\,dx + Q\,dy + R\,dz$$

be an exact differential of some function $\Phi(x, y, z)$ *is that*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \qquad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z},$$

for every point of the region. The function $\Phi(x, y, z)$ *is determined from the formula*

$$\Phi(x, y, z) = \int_{(a,b,c)}^{(x,y,z)} (P\,dx + Q\,dy + R\,dz) + \text{const.}$$

60. Applications of Line Integrals.

1. *Change of Variables in a Double Integral.* A formula for the change of variables in a double integral was developed in Sec. 46, and it is advisable to review Secs. 46 and 47 before reading this section. A more elegant, though less direct, derivation of the relation established there is given below with the aid of the concept of the line integral.

Consider again the equations of transformation of Sec. 46,

$$(60\text{-}1) \qquad x = \varphi_1(u, v), \qquad y = \varphi_2(u, v),$$

where $\varphi_1(u, v)$ and $\varphi_2(u, v)$ together with their first partial derivatives are continuous in the region under consideration. The equations (60-1) may be regarded as the equations of transformation from one set of rectangular cartesian axes u, v to another set of rectangular axes x, y. The area A of the region R enclosed by the curve C is given by the formula (54-4),

$$A = \int_C x\,dy.$$

Substituting for x and dy from (60-1) gives

$$A = \int_{C'} \varphi_1(u, v)\,d\varphi_2(u, v), \tag{60-2}$$

where C' is the boundary of the region R' in the uv-plane, which corresponds to the region R in the xy-plane.
But

$$d\varphi_2(u, v) = \frac{\partial \varphi_2}{\partial u}\,du + \frac{\partial \varphi_2}{\partial v}\,dv,$$

so that the integral (60-2) becomes

$$A = \int_{C'} \left[\varphi_1(u, v) \frac{\partial \varphi_2}{\partial u}\,du + \varphi_1(u, v) \frac{\partial \varphi_2}{\partial v}\,dv \right].$$

Setting

$$M(u, v) = \varphi_1 \frac{\partial \varphi_2}{\partial u},$$

and

$$N(u, v) = \varphi_1 \frac{\partial \varphi_2}{\partial v},$$

and applying Green's theorem (Sec. 55) gives

$$\begin{aligned} A &= \iint_{R'} \left| \frac{\partial}{\partial v}\left(\varphi_1 \frac{\partial \varphi_2}{\partial u} \right) - \frac{\partial}{\partial u}\left(\varphi_1 \frac{\partial \varphi_2}{\partial v} \right) \right| du\,dv \\ &= \iint_{R'} \left| \frac{\partial \varphi_1}{\partial v} \frac{\partial \varphi_2}{\partial u} - \frac{\partial \varphi_1}{\partial u} \frac{\partial \varphi_2}{\partial v} \right| du\,dv \\ &= \iint_{R'} |J(u, v)|\,du\,dv, \end{aligned}$$

where

$$J(u, v) = \begin{vmatrix} \dfrac{\partial\varphi_1}{\partial u} & \dfrac{\partial\varphi_2}{\partial u} \\ \dfrac{\partial\varphi_1}{\partial v} & \dfrac{\partial\varphi_2}{\partial v} \end{vmatrix}.$$

It follows from the mean-value theorem for double integrals (see problem in Sec. 42) that

$$dA = |J(\xi, \eta)|\, du\, dv.$$

The absolute value bars were introduced in the foregoing because the area is essentially positive.

2. *Work.* It will be assumed that a force $F(x, y)$ acts at every point of the xy-plane (Fig. 63). This force varies from point to point in magnitude and direction. An example of such conditions is the case of an electric field of force. The problem is to determine the work done on a particle moving from the point $A(a, b)$ to the point $B(c, d)$ along some curve C. Divide the arc AB of C into n segments by the points $P_1, P_2, \ldots, P_{n-1}$, and let $\Delta s_i = P_iP_{i+1}$. Then the force acting at P_i is $F(x_i, y_i)$. Let it be directed along the line P_iS, and let P_iT be the tangent to C at P_i, making an angle θ_i with P_iS.

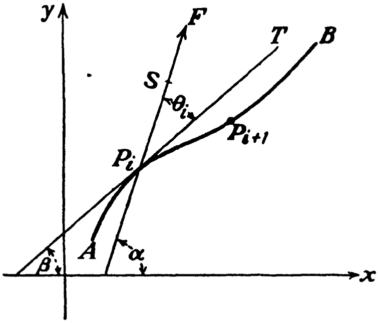

Fig. 63.

The component of force $F(x_i, y_i)$ along P_iT is $F \cos \theta_i$ and the element of work done on the particle in moving through the distance Δs_i is approximately $F(x_i, y_i) \cos \theta_i \, \Delta s_i$. The smaller Δs_i, the better this approximation will be. Therefore, the work done in moving the particle from A to B along C is

$$W = \lim_{n\to\infty} \sum_{i=0}^{n-1} F(x_i, y_i) \cos \theta_i \, \Delta s_i \equiv \int_C F(x, y) \cos \theta \, ds.$$

If α is the inclination of P_iS and β is the inclination of P_iT, then

$$\theta = \alpha - \beta \qquad \text{and} \qquad \cos\theta = \cos\alpha\cos\beta + \sin\alpha\sin\beta,$$

so that

$$W = \int_C F(x, y)(\cos \alpha \cos \beta + \sin \alpha \sin \beta)\, ds. \tag{60-3}$$

From the definition of α it is evident that

$$\begin{aligned} F \cos \alpha &= x\text{-component of } F \equiv X, \\ F \sin \alpha &= y\text{-component of } F \equiv Y. \end{aligned}$$

Moreover, since $\dfrac{dx}{ds} = \cos \beta$ and $\dfrac{dy}{ds} = \sin \beta$,

$$\cos \beta\, ds = dx \quad \text{and} \quad \sin \beta\, ds = dy.$$

Therefore (60-3) becomes

$$W = \int_C (X\, dx + Y\, dy),$$

which is a line integral of the form (53-3).

If C is a space curve, then an argument in every respect similar to the foregoing shows that the work done in producing a displacement along a curve C in a field of force where the components along the coordinate axes are X, Y, and Z is

$$W = \int_C (X\, dx + Y\, dy + Z\, dz).$$

To illustrate the use of this formula, the work done in displacing a particle of mass m along some curve C, joining the points A and B, will be calculated. It will be assumed that the particle is moving under the Newtonian law of attraction

$$F = \frac{km}{r^2},$$

where k is the gravitational constant and r is the distance from the center of attraction O (containing a unit mass) to a position of the particle (Fig. 64).

The component of force in the direction of the positive x-axis is

$$X = F \cos (x, r) = -\frac{km}{r^2} \cdot \frac{x}{r}.$$

Similarly,

$$Y = -\frac{km}{r^2} \cdot \frac{y}{r} \quad \text{and} \quad Z = -\frac{km}{r^2} \cdot \frac{z}{r}.$$

The work done in displacing the particle from A to B is

$$W = -\int_A^B \frac{km}{r^3}(x\,dx + y\,dy + z\,dz).$$

But

$$r = \sqrt{x^2 + y^2 + z^2} \qquad \text{and} \qquad dr = \frac{x\,dx + y\,dy + z\,dz}{r}.$$

Therefore,

$$W = -km\int_A^B \frac{dr}{r^2} = km\left[\frac{1}{r}\right]_A^B,$$

which depends only on the coordinates of the points A and B

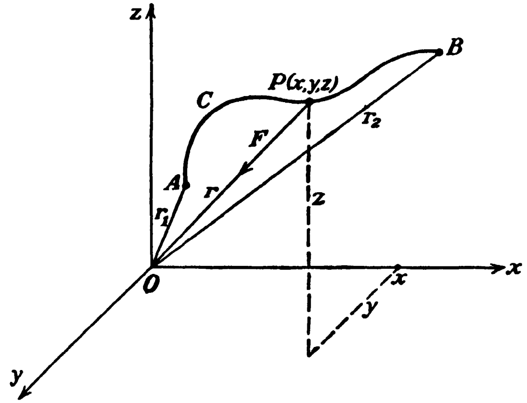

Fig. 64.

and not on the path C. Denoting the distances from O to A and B by r_1 and r_2, respectively, gives

$$W = km\left(\frac{1}{r_2} - \frac{1}{r_1}\right).$$

The quantity $\Phi \equiv \dfrac{km}{r}$ is known as the gravitational potential of the mass m. It is easily checked that

$$X = \frac{\partial\Phi}{\partial x}, \qquad Y = \frac{\partial\Phi}{\partial y}, \qquad Z = \frac{\partial\Phi}{\partial z},$$

so that the partial derivatives of the potential function Φ give the components of force along the coordinate axes. Moreover, the directional derivative of Φ in any direction s is

$$\begin{aligned}\frac{d\Phi}{ds} &= \frac{\partial\Phi}{\partial x}\frac{dx}{ds} + \frac{\partial\Phi}{\partial y}\frac{dy}{ds} + \frac{\partial\Phi}{\partial z}\frac{dz}{ds}\\ &= X\cos(x, s) + Y\cos(y, s) + Z\cos(z, s)\\ &= F_s,\end{aligned}$$

where F_s is the component of force in the direction s.

A conservative field of force is defined as a field of force in which the work done in producing a displacement between two fixed points is independent of the path. It is clear that in a conservative field the integral

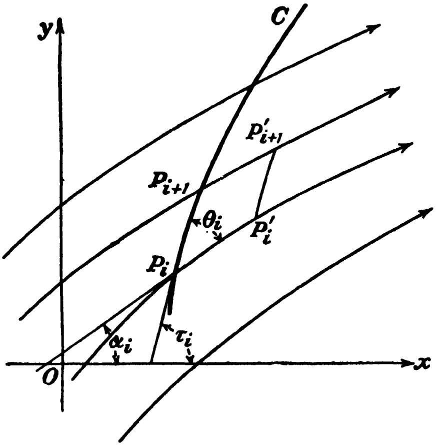

Fig. 65.

$$\int_C (X\,dx + Y\,dy + Z\,dz)$$

along every closed path is zero, so that the integrand is an exact differential.

3. *Flow of a Liquid.* Let C be a curve on a plane surface across which a liquid is flowing. The xy-plane will be chosen to coincide with the surface. The lines of flow are indicated in Fig. 65 by the curved arrows. It will be assumed that the flow of the liquid takes place in planes parallel to the xy-plane, and that the depth of the liquid is unity. The problem is to determine the amount of liquid that flows across C in a unit of time.

If v_i is the velocity of the liquid and α_i is the inclination of the tangent to the line of flow at P_i, then $v_x|_i = v_i \cos \alpha_i$ is the x-component of v_i and $v_y|_i = v_i \sin \alpha_i$ is the y-component of v_i. Let Δs_i denote the segment P_iP_{i+1} of C. A particle at P_i will move in time Δt to P_i' while a particle at P_{i+1} will move to P_{i+1}'. Therefore, the amount of liquid crossing P_iP_{i+1} in time Δt is equal to the volume of the cylinder whose altitude is unity and whose base is $P_iP_{i+1}P_{i+1}'P_i'$. Aside from infinitesimals of higher order this volume is $\Delta V_i = P_iP_i' \cdot P_iP_{i+1} \sin \theta_i$, in which θ_i denotes the angle between P_iP_i' and P_iP_{i+1}. But $P_iP_{i+1} = \Delta s_i$ and, except for infinitesimals of higher order, $P_iP_i' = v_i\,\Delta t$. Therefore $\Delta V_i = v_i\,\Delta t \cdot \Delta s_i \sin \theta_i$. The volume of liquid crossing C in a unit of time is

$$V = \lim_{n \to \infty} \sum_{i=1}^{n} v_i \sin \theta_i \,\Delta s_i.$$

If τ_i denotes the inclination of the tangent to C at P_i, then $\tau_i = \theta_i + \alpha_i$. Therefore,

$$\begin{aligned} v_i \sin \theta_i \,\Delta s_i &= v_i(\sin \tau_i \cos \alpha_i - \cos \tau_i \sin \alpha_i)\,\Delta s_i \\ &= v_i \cos \alpha_i \sin \tau_i \,\Delta s_i - v_i \sin \alpha_i \cos \tau_i \,\Delta s_i \\ &= v_x|_i\,\Delta y_i - v_y|_i\,\Delta x_i. \end{aligned}$$

Hence,

(60-4) $$V = \int_C (-v_y\,dx + v_x\,dy)$$

is the line integral which gives the amount of liquid that crosses C in a unit of time.

If the contour C is a closed one and the liquid is incompressible, then the net amount of liquid crossing C is zero, since as much liquid enters the region as leaves it. This assumes, of course, that the interior of C contains no sources or sinks. Thus, a steady flow of incompressible liquid is characterized by the equation

$$\int_C (-v_y\,dx + v_x\,dy) = 0,$$

over any closed contour not containing sources or sinks. This implies that (see Sec. 56)

(60-5) $$-\frac{\partial v_y}{\partial y} = \frac{\partial v_x}{\partial x},$$

which is an important equation of hydrodynamics known as the *equation of continuity*. Moreover, from Theorem 3, Sec. 56, it is known that there exists a function Ψ such that

$$\frac{\partial \Psi}{\partial x} = -v_y \quad \text{and} \quad \frac{\partial \Psi}{\partial y} = v_x. \tag{60-6}$$

This function Ψ is called the *stream function*, and it has a simple physical meaning, since

$$\Psi(x, y) = \int_{(a,b)}^{(x,y)} (-v_y\, dx + v_x\, dy)$$

represents the amount of liquid crossing, per unit time, any curve joining (a, b) with (x, y).

The function defined by the integral

$$\Phi(x, y) = \int_{(a,b)}^{(x,y)} (v_x\, dx + v_y\, dy), \tag{60-7}$$

is called the *velocity potential*. It is readily shown that

$$\frac{\partial \Phi}{\partial x} = v_x \quad \text{and} \quad \frac{\partial \Phi}{\partial y} = v_y. \tag{60-8}$$

Comparing (60-6) with (60-8), it is seen that

$$\frac{\partial \Phi}{\partial x} = \frac{\partial \Psi}{\partial y} \quad \text{and} \quad \frac{\partial \Phi}{\partial y} = -\frac{\partial \Psi}{\partial x}.$$

These are the celebrated Cauchy-Riemann differential equations.

If the integral (60-4) around a closed curve C does not vanish, then the region bounded by C may contain sources (if V is positive) or sinks (if V is negative). The presence of sources or sinks is characterized by the singularities of the function Ψ, that is, those points for which Ψ is not continuous or where its derivatives may cease to be continuous.*

The foregoing discussion is readily generalized to a steady flow of liquids in space. Instead of the integral (60-7) one will have

$$\Phi(x, y, z) = \int_{(a,b,c)}^{(x,y,z)} (v_x\, dx + v_y\, dy + v_z\, dz),$$

---

* See in this connection Sec. 57.

and if this integral is independent of the path C, the equations corresponding to (60-5) are

$$\frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x} = 0, \qquad \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} = 0, \qquad \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} = 0.$$

In such a case the integrand is an exact differential, and the velocity potential $\Phi(x, y, z)$ gives

$$\frac{\partial \Phi}{\partial x} = v_x, \qquad \frac{\partial \Phi}{\partial y} = v_y, \qquad \frac{\partial \Phi}{\partial z} = v_z.$$

4. *Thermodynamics.* If p, v, and U are the pressure, volume, and internal energy of a gas enclosed in a receptacle, then they are connected by the equation

$$f(p, v, U) = 0. \tag{60-9}$$

Hence, the state of a gas is completely determined from (60-9) by specifying any two of the three quantities, say p and v, which can be regarded as representing the coordinates of a point P in the pv-plane. If the state of the gas changes, the point P describes a curve C, and if the process is cyclic, the curve C will be a closed one.

It is important to know the amount Q of heat lost or absorbed by the gas while the gas in the receptacle (for example, steam in an engine cylinder) changes its state. Let Δp, Δv, and ΔU be the increments of pressure, volume, and energy when the amount ΔQ of heat is added. Then

$$\Delta Q = \Delta U + p\Delta v.$$

If p and v are considered the independent variables, then U is determined from (60-9), giving $U = U(p, v)$. Hence, approximately,

$$\Delta U = \frac{\partial U}{\partial p} \Delta p + \frac{\partial U}{\partial v} \Delta v.$$

Therefore, except for infinitesimals of higher order,

$$\Delta Q = \frac{\partial U}{\partial p} \Delta p + \frac{\partial U}{\partial v} \Delta v + p \,\Delta v = \frac{\partial U}{\partial p} \Delta p + \left(\frac{\partial U}{\partial v} + p\right) \Delta v.$$

As the state of the gas changes, p and v vary along a curve C in the pv-plane, and it is clear that the total amount of heat introduced into the gas during this change is

$$Q = \int_C \left[\frac{\partial U}{\partial p}\, dp + \left(\frac{\partial U}{\partial v} + p \right) dv \right]. \tag{60-10}$$

The integral (60-10) is obviously of the form

$$\int_C [M(p, v)\, dp + N(p, v)\, dv].$$

If, during the expansion of the gas, there is no gain or loss of heat, then the process is called *adiabatic.* It is characterized by the equation $dQ = 0$.

PROBLEM

Show with the aid of (60-5) that the velocity potential Φ satisfies the equation of Laplace,

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0.$$

CHAPTER VII

INFINITE SERIES

61. Infinite Series. Let $u_1, u_2, u_3, \cdots, u_n, \cdots$ be any sequence of quantities; then the symbol

$$\sum_{n=1}^{\infty} u_n \equiv u_1 + u_2 + u_3 + \cdots + u_n + \cdots \tag{61-1}$$

is called an *infinite series.* With the aid of the symbol (61-1) one can form a new sequence $\{s_i\}$ whose elements are sums of a finite number of the terms of the series (61-1):

$$\begin{aligned} s_1 &= u_1, \\ s_2 &= u_1 + u_2, \\ s_3 &= u_1 + u_2 + u_3, \\ &\cdots\cdots\cdots\cdots\cdots, \\ s_n &= u_1 + u_2 + u_3 + \cdots + u_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

The numbers s_i will be called the ith *partial sums* of the series (61-1).

Definition. *An infinite series* $\sum_{n=1}^{\infty} u_n$ *is said to converge if the sequence of partial sums* $\{s_i\}$ *is convergent. A series which does not converge is called divergent.*

Stated in symbolic form this definition of the convergence of an infinite series is*

$$\lim_{n \to \infty} s_n = S, \qquad \text{or} \qquad |S - s_n| < \epsilon, \qquad \text{for all} \qquad n \geq p. \tag{61-2}$$

The number S is frequently called the *sum of the series* (61-1), although one should clearly realize that it represents the limit of the partial sums s_n.

---

* See Sec. 3.

A necessary and sufficient condition for convergence of any sequence was given in Sec. 5, and it will be restated here for easy reference.

A necessary and sufficient condition for the convergence of $\sum_{n=1}^{\infty} u_n$ *is that for any* $\epsilon > 0$, *one can find a positive integer* p *(depending on* ϵ*) such that*

$$|s_n - s_m| < \epsilon$$

whenever n *and* m *both exceed* p.

It is clear that the series cannot converge if its partial sums s_n are unbounded, that is, if $|s_n|$ can be made greater than any preassigned positive number M. However, the boundedness of partial sums does not ensure the convergence, as can be seen from the simple example of the series

$$1 - 1 + 1 - 1 + \cdots.$$

Here $s_1 = 1$, $s_2 = 0$, $s_3 = 1$, $\cdots$, $s_{2n-1} = 1$, $s_{2n} = 0$, $\cdots$, and the s_n are clearly bounded, but the sequence

$$1, 0, 1, 0, 1, \cdots$$

does not converge.

Infinite series possess some properties which follow directly from the definition of convergence and from the fundamental criterion of convergence of sequences:

(*A*) *The convergence or divergence of a series* $\sum_{n=1}^{\infty} u_n$ *is not destroyed by adding to or subtracting from it a finite number of terms.*

This follows directly from the fundamental criterion. Let the series be convergent; then for any $\epsilon > 0$ one can find a positive integer p such that for any pair of numbers $m > p$ and $n > p$,

$$|s_n - s_m| < \epsilon.$$

This condition, certainly, remains unaltered by the addition (or subtraction) of a finite number of terms. To be sure, the sum of the series will be altered by the addition (or subtraction) of a finite number of terms, but the property of convergence will not be affected. Thus, if the sum of the series $\sum_{n=1}^{\infty} u_n$ is S, then the addition of a finite number of terms $a_1 + a_2 + \cdots + a_k$ will give for the sum of the new series $S + a_1 + a_2 + \cdots + a_k$. A similar argument can be supplied if the series diverges.

(*B*) *The limit of the general term* u_n *of the series* $\sum_{n=1}^{\infty} u_n$ *is necessarily zero if the series is convergent.*

For

$$u_n = s_n - s_{n-1},$$

and if the series converges, then

$$\lim_{n\to\infty} s_n = S \qquad \text{and} \qquad \lim_{n\to\infty} s_{n-1} = S.$$

Thus,

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = S - S = 0.$$

The condition enunciated is necessary but not sufficient, that is, a series may diverge even though $\lim_{n\to\infty} u_n = 0$. A classical example illustrating this case is the harmonic series

✓ $$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots,$$

in which s_n increases without limit as n increases, as will be shown next.

Since

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n+n} > n \cdot \frac{1}{2n} = \frac{1}{2},$$

it is possible, beginning with any term of the series, to add a definite number of terms and obtain a sum greater than $\frac{1}{2}$. If $n = 2$,

$$\tfrac{1}{3} + \tfrac{1}{4} > \tfrac{1}{2};$$

$n = 4$,

$$\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8} > \tfrac{1}{2};$$

$n = 8$,

$$\tfrac{1}{9} + \tfrac{1}{10} + \cdots + \tfrac{1}{16} > \tfrac{1}{2};$$

$n = 16,$

$$\tfrac{1}{17} + \tfrac{1}{18} + \cdots + \tfrac{1}{32} > \tfrac{1}{2}.$$

Thus it is possible to group the terms of the harmonic series*

$$1 + \tfrac{1}{2} + (\tfrac{1}{3} + \tfrac{1}{4}) + (\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}) + \cdots$$

in such a way that the sum of the terms in each parenthesis exceeds $\tfrac{1}{2}$, and, since the series

$$1 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \cdots$$

is obviously divergent, the harmonic series is divergent also.

(*C*) *Let the sum of the series* $\sum_{n=1}^{\infty} u_n$ *be* S. *If a new series is formed by multiplying each term by a constant* k *to form the series*

$$ku_1 + ku_2 + \cdots + ku_n + \cdots,$$

then the sum of the latter series is kS.

For, let

$$s_n = u_1 + u_2 + \cdots + u_n$$

and

$$s_n' = ku_1 + ku_2 + \cdots + ku_n;$$

then

$$s_n' = ks_n.$$

Hence,

$$\lim_{n\to\infty} s_n' = k \lim_{n\to\infty} s_n = kS.$$

(*D*) *Given the two convergent series*

$$u_1 + u_2 + \cdots + u_n + \cdots$$

and

$$v_1 + v_2 + \cdots + v_n + \cdots.$$

---

* The fact that the associative law holds unrestrictedly for series of *positive* terms follows from the fact that the partial sums of the series

(1) $$(u_1 + u_2 + \cdots + u_r) + (u_{r+1} + \cdots + u_k) + \cdots$$

are subsequences of the partial sums of the series

(2) $$u_1 + u_2 + u_3 + \cdots + u_n + \cdots.$$

Hence, (1) and (2) will diverge or converge together.

Let the sum of the first series be U and that of the second be V, then the series

$$(61\text{-}3) \qquad (u_1 \pm v_1) + (u_2 \pm v_2) + \cdots + (u_n \pm v_n) + \cdots$$

has the sum $U \pm V$.

Let

$$s_n = u_1 + u_2 + \cdots + u_n,$$
$$s_n' = v_1 + v_2 + \cdots + v_n;$$

then

$$s_n \pm s_n' = (u_1 \pm v_1) + (u_2 \pm v_2) + \cdots + (u_n \pm v_n),$$

but this is precisely the nth partial sum of the series (61-3), which can be designated by s_n''.

Then

$$\lim_{n\to\infty} (s_n \pm s_n') = \lim_{n\to\infty} s_n''$$

or

$$U \pm V = \lim_{n\to\infty} s_n''.$$

This result can be stated as follows:

Convergent series may be added or subtracted term by term. It should be noted that this property has been established for a particular mode of addition. It will be seen below that the sum of the series may depend on the order in which the terms are added.

62. Series of Positive Terms. This section will be concerned exclusively with series whose terms are nonnegative numbers. The property (A) of the preceding section states that one can discard any finite number of terms without affecting the convergence of the series, so that the conclusions of this section, in regard to convergence, will also be applicable to any series containing a finite number of negative terms.

Theorem 1. *A series of positive terms is necessarily convergent if its partial sums are bounded.*

Since the terms of the series are positive, the partial sums will form a monotone-increasing sequence of numbers, that is,[*]

$$s_1 \leq s_2 \leq s_3 \leq \cdots \leq s_n \leq \cdots,$$

---

[*] The equality sign allows some terms to be zero.

and a bounded monotone sequence is necessarily convergent (see Sec. 6).

Theorem 2. *If* $\sum_{n=1}^{\infty} u_n$ *is a convergent series of positive terms, and* $\{a_n\}$ *is any bounded sequence of positive numbers, then the series* $\sum_{n=1}^{\infty} a_n u_n$ *is convergent.*

Since the numbers a_n are bounded,

$$a_n < M.$$

Then the nth partial sum of $\sum_{n=1}^{\infty} a_n u_n$ is less than Ms_n, where s_n is the nth partial sum of $\sum_{n=1}^{\infty} u_n$. From Theorem 1 it follows that $\sum_{n=1}^{\infty} a_n u_n$ is convergent, since its partial sums are bounded.

Example 1. Consider

$$\sum_{n=1}^{\infty} a^n \equiv a + a^2 + a^3 + \cdots + a^n + \cdots .$$

(*a*) If $a \geq 1$, $s_n \geq n$ so that the partial sums are not bounded, and hence, the series diverges.

(*b*) If $a < 1$, the series is a convergent geometric series. For the sum of n terms of a geometric progression of ratio a is*

$$s_n = \frac{a(1 - a^n)}{1 - a} = \frac{a}{1 - a} - \frac{a^{n+1}}{1 - a}.$$

Thus,

$$s_n < \frac{a}{1 - a},$$

---

* The sum of n terms of a geometrical progression of ratio r and whose first term is a is $s_n = \dfrac{a(1 - r^n)}{1 - r}$.

so that the partial sums are bounded. Consequently the series converges. In fact, $\lim_{n\to\infty} s_n = \frac{a}{1-a}$, since $\lim_{n\to\infty} a^{n+1} = 0$, if $a < 1$.

Example 2. Consider

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{n(n+1)}+\cdots.$$

This series can be written as

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{n}-\frac{1}{n+1}\right)+\cdots.$$

Hence, the nth partial sum is

$$s_n = 1 - \frac{1}{n+1}.$$

Thus, s_n is bounded, so that the series is convergent.

Theorem 3 (Comparison Test). *Let* $\sum_{n=1}^{\infty} u_n$ *and* $\sum_{n=1}^{\infty} v_n$ *be two series of positive terms, the first of which is known to be convergent and the second divergent.*

(*a*) *If the terms of a given series* $\sum_{n=1}^{\infty} a_n$ *of positive terms are such that* $a_n \leq u_n$, *for every* $n \geq m$, *then the series* $\sum_{n=1}^{\infty} a_n$ *is convergent.*

(*b*) *If the terms of a given series* $\sum_{n=1}^{\infty} a_n$ *of positive terms are such that* $a_n \geq v_n$, *for every* $n \geq m$, *then the series* $\sum_{n=1}^{\infty} a_n$ *is divergent.*

In order to establish part (*a*) of the theorem, assume* that the inequality $a_n \leq u_n$ is satisfied for every term of the series, so that $m = 1$. Denote the nth partial sum of the given series $\sum_{n=1}^{\infty} a_n$ by s_n, and that of the series $\sum_{n=1}^{\infty} u_n$ by U_n. Then

$$s_n \leq U_n.$$

---

* This entails no loss of generality. See property (*A*). Sec. 61.

Since the sequences $\{s_n\}$ and $\{U_n\}$ are monotone increasing and the series $\sum_{n=1}^{\infty} u_n$ is convergent, it follows that

$$s_n \leq \lim_{n \to \infty} U_n = U.$$

Thus, the sequence $\{s_n\}$ is bounded, and, therefore, the series $\sum_{n=1}^{\infty} a_n$ is convergent.

In order to prove the second part of the theorem, denote the nth partial sum of $\sum_{n=1}^{\infty} v_n$ by V_n; then

$$s_n \geq V_n.$$

But, by hypothesis $\sum_{n=1}^{\infty} v_n$ is divergent, so that $\{V_n\}$ is an unbounded sequence, and hence $\sum_{n=1}^{\infty} a_n$ diverges.

This test was used essentially in establishing the divergence of the harmonic series in Sec 61.

Theorem 4 **(Cauchy's Root Test).** *If a given series* $\sum_{n=1}^{\infty} a_n$ *of positive terms is such that*

(*a*) $\sqrt[n]{a_n} \leq r < 1$, *for every* $n \geq m$, *where* r *is independent of* n, *then the series* $\sum_{n=1}^{\infty} a_n$ *is convergent.*

If, however,

(*b*) $\sqrt[n]{a_n} \geq 1$, *for every* $n \geq m$, *then the series* $\sum_{n=1}^{\infty} a_n$ *is divergent.*

There is no loss of generality in assuming that the inequalities (*a*) and (*b*) are satisfied for $m = 1$. It follows from (*a*) that

$$a_n \leq r^n,$$

so that no term of the given series exceeds the corresponding term of the geometric series

$$r + r^2 + \cdots + r^n + \cdots,$$

which is convergent, since $r < 1$. Thus, by Theorem 3, the given series converges.

The truth of the assertion of the second part of the theorem follows from property (B) of Sec. 61. If $\sqrt[n]{a_n} \geq 1$, the general term of the given series does not tend to zero, and hence, the series must diverge.

Corollary. *If the given series* $\sum_{n=1}^{\infty} a_n$ *is such that*

$$\lim_{n\to\infty} \sqrt[n]{a_n} = r < 1,$$

then the series is convergent.

If

$$\lim_{n\to\infty} \sqrt[n]{a_n} > 1,$$

then the series $\sum_{n=1}^{\infty} a_n$ *diverges.*

Consider first the case when $r < 1$, and choose a positive number ϵ so small that

$$r + \epsilon < 1.$$

Then, for any sufficiently large n, the quantity $\sqrt[n]{a_n}$ can be made to differ from r by less than ϵ, that is,

$$r - \epsilon < \sqrt[n]{a_n} < r + \epsilon.$$

But $r + \epsilon = r' < 1$, and it follows from Theorem 4 that the series converges.

The proof of the case when $r > 1$ follows directly from property (B) of Sec. 61.

Theorem 5 **(d'Alembert's Ratio Test*).** *If the terms of a series* $\sum_{n=1}^{\infty} a_n$ *of positive terms satisfy, from some value of n onward, the inequality*

$$\frac{a_{n+1}}{a_n} \leq r < 1,$$

---

* Also known as *Cauchy's ratio test.*

where r is independent of n, then the given series is convergent. If, however, from and after some value of n

$$\frac{a_{n+1}}{a_n} \geq 1,$$

then the given series diverges.

Let

$$\frac{a_{n+1}}{a_n} \leq r, \qquad (r < 1).$$

Then,

$$a_{n+1} \leqq ra_n, \tag{62-1}$$

and the inequality may be assumed to be satisfied for all $n \geq 1$. Hence,

$$\begin{aligned} a_{n+1} &\leq ra_n, \\ a_n &\leq ra_{n-1}, \\ a_{n-1} &\leq ra_{n-2}, \\ &\cdots\cdots, \\ a_3 &\leq ra_2, \\ a_2 &\leq ra_1. \end{aligned}$$

Forming the product of the left-hand members and the product of the right-hand members of these inequalities and dividing out common factors give

$$a_{n+1} \leq r^n a_1, \qquad (n = 1, 2, \cdots).$$

Thus, the terms of the given series are less than or equal to the corresponding terms of the geometric series

$$a_1(r + r^2 + \cdots + r^n + \cdots),$$

which is convergent since $r < 1$. Hence, the given series is convergent by the comparison test.

The proof of the second part of the theorem is left as an exercise for the reader.

Corollary. *If the given series $\sum_{n=1}^{\infty} a_n$ is such that*

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = r < 1,$$

then the given series converges.

If

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} > 1,$$

then the series $\sum_{n=1}^{\infty} a_n$ *diverges.*

The proof is identical with that of the corollary to Theorem 4, where $\frac{a_{n+1}}{a_n}$ replaces $\sqrt[n]{a_n}$.

In applying Theorems 4 and 5, to establish the convergence of the series $\sum_{n=1}^{\infty} a_n$, it is essential to note that $\sqrt[n]{a_n}$ and $\frac{a_{n+1}}{a_n}$ must be less than some fixed number r smaller than unity for all values from some n onward. For example, in the harmonic series

$$\sum_{n=1}^{\infty} a_n \equiv \sum_{n=1}^{\infty} \frac{1}{n},$$

each of the expressions

$$\frac{a_{n+1}}{a_n} = \frac{n}{n+1} \qquad \text{and} \qquad \sqrt[n]{a_n} = \sqrt[n]{\frac{1}{n}}$$

is always less than 1, but the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. The difficulty here lies in the fact that *no fixed number r, independent of n, can be assigned for which*

$$\frac{n}{n+1} \leq r < 1.$$

This ratio can be made to approach arbitrarily near unity, which is in violation of the requirement that

$$\frac{a_{n+1}}{a_n} \leq r < 1.$$

In other words, $\sqrt[n]{a_n}$ and $\frac{a_{n+1}}{a_n}$ must ultimately be less than some fixed number r, which is smaller than unity.

In general, the ratio test is easier to apply than the root test, but the latter is more powerful. It will be seen from the example given below that the root test is capable of furnishing positive information when the ratio test fails. It may be remarked that the root test is always applicable when the ratio test is applicable.*

The corollaries to Theorems 4 and 5 are of most frequent use in practical applications, but if the limits involved fail to exist, one may have to apply the theorems themselves. Thus, consider the series

$$r|\sin \alpha| + r^2|\sin 2\alpha| + \cdots + r^n|\sin n\alpha| + \cdots,$$

where α is an arbitrary constant.

In this case

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{r^{n+1}}{r^n}\left|\frac{\sin (n+1)\alpha}{\sin n\alpha}\right| \\ &= r\left|\frac{\sin (n+1)\alpha}{\sin n\alpha}\right|.\end{aligned}$$

This ratio does not approach any limit. In fact†

$$\left|\frac{\sin (n+1)\alpha}{\sin n\alpha}\right|$$

may assume values greater than unity as $n \to \infty$. But the root test gives

$$\sqrt[n]{a_n} = \sqrt[n]{r^n|\sin n\alpha|} = r\sqrt[n]{|\sin n\alpha|} \leq r.$$

Thus, if $r < 1$, the series converges, no matter what α is. It should be noted here that $\lim\limits_{n\to\infty} \sqrt[n]{a_n}$ does not exist.

As an example of the application of the corollaries to Theorems 4 and 5, consider the series

$$x + 2x^2 + 3x^3 + \cdots + nx^n + \cdots,$$

---

* For details, see K. Knopp, Theory and Application of Infinite Series, Sec. 36, English edition.

† However, it is impossible to find an integer m such that for all $n \geq m$, $\left|\frac{\sin (n+1)\alpha}{\sin n\alpha}\right|$ is greater than unity.

where $x \geq 0$. The test ratio is

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{n}\frac{x^{n+1}}{x^n} = \left(1 + \frac{1}{n}\right)x,$$

and

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = x.$$

Thus the series converges if $0 \leq x < 1$ and diverges if $x > 1$. When $x = 1$, the test fails, but the series

$$1 + 2 + 3 + \cdots + n + \cdots$$

is obviously divergent.

The root test furnishes the same information, since

$$\sqrt[n]{a_n} = \sqrt[n]{n}\, x.$$

But*

$$\lim_{n \to \infty} \sqrt[n]{n} = 1,$$

so that

$$\lim_{n \to \infty} \sqrt[n]{n}\, x = x,$$

and if $0 \leq x < 1$, then the series is convergent.

The series

$$\frac{1!}{10} + \frac{2!}{10^2} + \frac{3!}{10^3} + \cdots + \frac{n!}{10^n} + \cdots$$

is readily shown to be divergent by the application of the ratio test. For

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = \lim_{n \to \infty} \frac{(n+1)!10^n}{n!10^{n+1}} = \lim_{n \to \infty} \frac{n+1}{10} = \infty.$$

The ratio test, in the form given in the corollary to Theorem 5, is very easy to apply, and if it fails in a given case, one may attempt to find a suitable comparison series and apply Theorem 3.

---

* $\log \sqrt[n]{n} = \frac{1}{n} \log n$, which tends to zero when $n \to \infty$. Hence

$$\sqrt[n]{n} \to e^0 = 1.$$

As an illustration of this circumstance, consider

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots.$$

In this case the test ratio is

$$\frac{a_{n+1}}{a_n}=\frac{n^2}{(n+1)^2},$$

so that

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=1.$$

Also,

$$\sqrt[n]{a_n}=\sqrt[n]{\frac{1}{n^2}},\qquad\text{so that}\qquad\lim_{n\to\infty}\sqrt[n]{a_n}=1.$$

Hence, neither of the tests given in the corollaries furnishes information regarding convergence. In this case, one may resort to the comparison test.

Now consider the more general series,

$$1+\frac{1}{2^p}+\frac{1}{3^p}+\cdots+\frac{1}{n^p}+\cdots. \tag{62-2}$$

It will be shown that this series converges for every $p > 1$ and diverges whenever $0 < p \leq 1$.

The terms of the series (62-2) may be grouped as follows:

$$1+\left(\frac{1}{2^p}+\frac{1}{3^p}\right)+\left(\frac{1}{4^p}+\frac{1}{5^p}+\frac{1}{6^p}+\frac{1}{7^p}\right)+\cdots + \left(\frac{1}{(2^{k-1})^p}+\cdots+\frac{1}{(2^k-1)^p}\right)+\cdots, \tag{62-3}$$

where the kth group contains 2^{k-1} terms. The sum of the terms in each parenthesis is less than the number of terms in that parenthesis multiplied by the first term. Hence, the terms of the series

$$1+2\cdot\frac{1}{2^p}+4\cdot\frac{1}{4^p}+\cdots+2^{k-1}\frac{1}{(2^{k-1})^p}+\cdots \tag{62-4}$$

are not less than the corresponding terms of the series (62-3). But the series (62-4) is a geometric series of ratio $\frac{1}{2^{p-1}}$. Hence, if $p > 1$, the ratio is less than unity, so that (62-4), and therefore (62-2), will converge.

It remains to establish the divergence of the series for $p \leq 1$. Note that $p = 1$ gives

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots,$$

which is known to be divergent. But the nth term of this series is not greater than the corresponding term of the series (62-2) if $p < 1$, since

$$\frac{1}{n^p} \geq \frac{1}{n}.$$

Hence, by the comparison test, the series (62-2) diverges for $p < 1$.

The series (62-2) and the geometric series are frequently used as the standard comparison series.

A similar argument can be used to establish the divergence of the series

$$\text{(62-5)} \qquad \frac{1}{2 \log 2} + \frac{1}{3 \log 3} + \cdots + \frac{1}{n \log n} + \cdots.$$

Let the terms of the series be grouped as is indicated below:

$$\frac{1}{2 \log 2} + \left(\frac{1}{3 \log 3} + \frac{1}{4 \log 4}\right) + \left(\frac{1}{5 \log 5} + \cdots + \frac{1}{8 \log 8}\right) + \left(\frac{1}{9 \log 9} + \cdots + \frac{1}{16 \log 16}\right) + \cdots.$$

Evidently,

$$\frac{1}{3 \log 3} + \frac{1}{4 \log 4} > \frac{2}{4 \log 4} = \frac{1}{2}\frac{1}{2 \log 2},$$

$$\frac{1}{5 \log 5} + \cdots + \frac{1}{8 \log 8} > \frac{4}{8 \log 8} = \frac{1}{3}\frac{1}{2 \log 2},$$

$$\frac{1}{9 \log 9} + \cdots + \frac{1}{16 \log 16} > \frac{8}{16 \log 16} = \frac{1}{4}\frac{1}{2 \log 2},$$

. . . .

Therefore, the terms of the series

$$\text{(62-6)} \qquad \frac{1}{2 \log 2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots\right)$$

are not greater than the corresponding terms of the given series, and the series (62-6) is obviously divergent. It follows that the series (62-5) is divergent.

As a further illustration, let it be required to test the series

$$\frac{1}{1\cdot 2}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{(2n-1)2n}+\cdots.$$

Here,

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{1}{(2n+1)(2n+2)}\cdot\frac{(2n-1)2n}{1}$$

$$=\lim_{n\to\infty}\frac{4n^2-2n}{4n^2+6n+2}=\lim_{n\to\infty}\frac{1-\dfrac{1}{2n}}{1+\dfrac{3}{2n}+\dfrac{1}{2n^2}}=1.$$

Hence, the test fails, but if the given series be compared with the series (62-2), for $p = 2$, it is seen to be convergent.

PROBLEMS

1. Prove that $\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\cdots+\frac{1}{2n}+\cdots$ diverges.

2. Prove that $1+\frac{1}{3}+\frac{1}{5}+\cdots+\frac{1}{2n-1}+\cdots$ diverges.

3. Test the series $1+\frac{2}{2^1}+\frac{3}{2^2}+\cdots+\frac{n}{2^{n-1}}+\cdots$.

4. Test the series $1+\frac{1}{\log 2}+\frac{1}{\log 3}+\cdots+\frac{1}{\log n}+\cdots$.

5. Test for convergence:

(*a*) $\frac{1}{2}+\frac{1}{2\cdot 2^2}+\frac{1}{3\cdot 2^3}+\frac{1}{4\cdot 2^4}+\cdots$;

(*b*) $\frac{1}{3}+\frac{1\cdot 2}{3\cdot 5}+\frac{1\cdot 2\cdot 3}{3\cdot 5\cdot 7}+\cdots$;

(*c*) $1+\frac{2!}{2^2}+\frac{3!}{3^2}+\cdots$;

(*d*) $\frac{1}{2}+\frac{2}{2^2}+\frac{3}{2^3}+\frac{4}{2^4}+\cdots$.

6. Show that $1+\frac{1}{\sqrt{2\cdot 3}}+\cdots+\frac{1}{\sqrt{n(n+1)}}+\cdots$ diverges.

7. Show that the series $1 + \frac{1}{5} + \frac{1}{10} + \cdots + \frac{1}{1 + n^2} + \cdots$ is convergent.

8. Show that $\sum_{n=2}^{\infty} \frac{1}{(\log n)^n}$ converges.

9. Show that $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n(1 + n^2)}}$ converges.

10. Show that $\sum_{n=1}^{\infty} \frac{n!}{n^2}$ diverges.

11. Show that $\sum_{n=2}^{\infty} \frac{1}{\sqrt{n} \log n}$ diverges.

12. Show that $\sum_{n=1}^{\infty} \frac{x^n}{n^n}$ converges, if $x \geq 0$.

13. Show that $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converges, if $x \geq 0$.

14. Show that $\sum_{n=1}^{\infty} \frac{x^n}{n}$ converges, if $0 \leq x < 1$.

15. Show that $\sum_{n=1}^{\infty} n!x^n$ diverges, if $x > 0$.

63. More General Tests. The tests established in Sec. 62 are those most commonly used in practice. There are many more general criteria of convergence than those given in the preceding section, but their application is considerably more involved.* Some of the sharper criteria are discussed in this section.

Theorem I (Cauchy's Integral Test). *Let $\sum_{n=1}^{\infty} a_n$ be a given series of decreasing positive terms. If there exists a positive monotone*

---

* See, for example, KNOPP, K., Theory and Application of Infinite Series.

decreasing function $f(x)$ for $x \geq 1$, such that $f(n) = a_n$, then the given series converges if the integral

$$\int_1^{\infty} f(x)\,dx$$

exists; the given series diverges if the integral does not exist.

By hypothesis $f(n) = a_n$, therefore, the nth partial sum of the given series is

$$s_n = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} f(i).$$

Let the values of $f(i)$ be plotted as the ordinates (Fig. 66); then each term of the series can be interpreted as the area of a rectangle of base unity and height $f(i)$. Each of these rectangles extends above the curve $y = f(x)$, so that the area of the sum of the n rectangles is greater than the area under the locus of $y = f(x)$ bounded by the x-axis, and the lines $x = 1$ and $x = n + 1$. Thus,

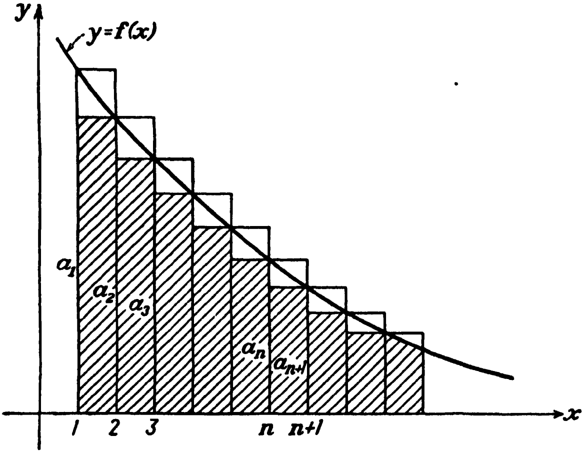

FIG. 66.

(63-1)
$$s_n > \int_1^{n+1} f(x)\,dx. > \int_1^{\infty} f(x)\,dx$$

hence if $\int_1^{\infty} f(x)dx$ diverges diverge

On the other hand, the sum of the areas of the inscribed rectangles,

$$a_2 + a_3 + \cdots + a_{n+1} = \sum_{n=2}^{n+1} f(n),$$

is less than the value of the integral $\int_1^{n+1} f(x)\,dx$, so that one can write

$$\sum_{n=2}^{n+1} f(n) < \int_1^{n+1} f(x)\,dx.$$

But

$$\sum_{n=2}^{n+1} f(n) = s_{n+1} - a_1.$$

Hence, the inequality just above can be written as

$$(63\text{-}2) \qquad < S_n < s_{n+1} < \int_1^{n+1} f(x)\,dx + a_1.$$

But the function $y = f(x)$ is positive; hence

$$\int_1^{n+1} f(x)\,dx < \int_1^{\infty} f(x)\,dx.$$

If the integral on the right converges to a value L, then it follows from the inequality (63-2) that

$$s_{n+1} < L + a_1.$$

Thus, the partial sums are bounded, and by Theorem 1, Sec. 62, the series $\sum_{n=1}^{\infty} a_n$ converges if the integral does. If, on the other hand, the integral becomes infinite, it follows from the inequality (63-1) that the series diverges.

It may be remarked that it is not necessary to demand that $f(n)$ be equal to a_n for all values of n beginning with $n = 1$. If m is any number, greater than unity, and if the condition $f(n) = a_n$ is satisfied for $n \geq m$, then the series $\sum_{n=1}^{\infty} a_n$ and the integral $\int_m^{\infty} f(x)\,dx$ will converge and diverge together.

Example 1. Test the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots.$$

Here $f(x) = \frac{1}{x}$, since $f(n) = \frac{1}{n}$. Now the integral

$$\int_1^n \frac{1}{x}\,dx = \log n$$

and

$$\int_1^\infty \frac{dx}{x} \equiv \lim_{n\to\infty} \int_1^n \frac{dx}{x} = \lim_{n\to\infty} \log n = \infty,$$

so the given series is divergent.

Example 2. Test

$$\frac{1}{2} + \frac{1}{5} + \frac{1}{10} + \cdots + \frac{1}{n^2+1} + \cdots.$$

Here

$$f(x) = \frac{1}{1+x^2}$$

and

$$\int_1^n \frac{dx}{1+x^2} = \tan^{-1} n - \tan^{-1} 1,$$

$$\lim_{n\to\infty} (\tan^{-1} n - \tan^{-1} 1) = \frac{\pi}{4},$$

and, since the integral converges, the given series is convergent.

The test of d'Alembert (Sec. 62) is merely a particular case of an extremely general test due to E. Kummer, from which a galaxy of important special tests can be deduced.

Theorem 2 (Kummer's Test). *The series of positive terms* $\sum_{n=1}^{\infty} u_n$ *is convergent if there exists a sequence of positive numbers,* $\{a_n\}$, *such that from some value of n onward,*

$$(63\text{-}3) \qquad a_n \frac{u_n}{u_{n+1}} - a_{n+1} \geq r > 0.$$

However, if

$$(63\text{-}4) \qquad a_n \frac{u_n}{u_{n+1}} - a_{n+1} \leq 0,$$

then the given series diverges, provided that the series $\sum_{n=1}^{\infty} \frac{1}{a_n}$ *diverges.*

It may be assumed, without loss of generality, that (63-3) is satisfied for $n \geqq 1$.

Setting $n = 1, 2, 3, 4, \cdots, n - 1$ in (63-3) gives,

$$\begin{aligned} a_1u_1 - a_2u_2 &\geq ru_2, \\ a_2u_2 - a_3u_3 &\geq ru_3, \\ \cdots\cdots&\cdots\cdots, \\ a_{n-1}u_{n-1} - a_nu_n &\geq ru_n. \end{aligned}$$

Adding these inequalities term by term gives

$$a_1u_1 - a_nu_n \geq r(s_n - u_1), \tag{63-5}$$

where s_n is the sum of the first n terms of the given series. It follows from (63-5) that

$$rs_n \leqq a_1u_1 - a_nu_n + ru_1,$$

or

$$\begin{aligned} s_n &\leqq \frac{(a_1 + r)u_1}{r} - \frac{a_nu_n}{r} \\ &< \frac{(a_1 + r)u_1}{r}. \end{aligned}$$

Thus the sum of any number of terms of the given series is bounded $\left(\text{since } \frac{(a_1 + r)u_1}{r} \text{ is independent of } n\right)$, and, hence, the series converges by Theorem 1, Sec. 62.

In order to establish the second part of the theorem, note that (63-4) gives

$$a_1u_1 \leqq a_2u_2 \leq \cdots \leq a_nu_n.$$

Hence,

$$u_n \geq a_1u_1 \cdot \frac{1}{a_n},$$

and

$$\sum_{n=1}^{\infty} u_n \geq a_1u_1 \sum_{n=1}^{\infty} \frac{1}{a_n}.$$

By hypothesis, the series $\sum_{n=1}^{\infty} \frac{1}{a_n}$ diverges; consequently, the series $\sum_{n=1}^{\infty} u_n$ also diverges.

It will be observed that for $a_n = 1$, $(n = 1, 2, 3, \cdots)$, the test of Kummer specializes to that of d'Alembert.*

If the series $\sum_{n=1}^{\infty} u_n$ is such that $\lim_{n \to \infty} \frac{u_{n+1}}{u_n} = 1$, the ratio test fails. A test due to Gauss frequently enables one to determine the behavior of the series in this doubtful case.

Theorem 3 (Gauss' Test). *If the series* $\sum_{n=1}^{\infty} u_n$ *of positive terms is such that the ratio* $\frac{u_n}{u_{n+1}}$ *can be written in the form*

$$\frac{u_n}{u_{n+1}} = 1 + \frac{h}{n} + \frac{A(n)}{n^2},$$

where $A(n)$ *is a bounded function of* n *as* $n \to \infty$, *then the series converges if* $h > 1$, *and diverges if* $h \leq 1$.

The proof of this test follows from the test of Kummer. Consider first the case when $h \neq 1$, and set a_n in Kummer's test equal to n. Then

$$\begin{aligned}\lim_{n \to \infty} \left(a_n \frac{u_n}{u_{n+1}} - a_{n+1}\right) &= \lim_{n \to \infty} \left[n\left(1 + \frac{h}{n} + \frac{A(n)}{n^2}\right) - n - 1\right] \\ &= h - 1.\end{aligned}$$

If $h > 1$, then $h - 1 = r > 0$, and it follows from Kummer's test that the given series is convergent. On the other hand, if $h - 1 < 0$, the test of Kummer ensures divergence, since $\sum_{n=1}^{\infty} \frac{1}{n}$ is a divergent series. It remains to consider the case when $h = 1$.

Note that the series $\sum_{n=1}^{\infty} \frac{1}{n \log n}$ is divergent,† and set a_n in Kummer's test equal to $n \log n$. Then,

$$a_n \frac{u_n}{u_{n+1}} - a_{n+1} = n \log n\left(1 + \frac{1}{n} + \frac{A(n)}{n^2}\right) - (n + 1) \log (n + 1)$$

---

* See also Raabe's test in Prob. 6, p. 233.

† This fact can be readily established with the aid of the integral test of Cauchy; see also (62-5).

$$= n \log n + \log n + \frac{A(n) \log n}{n} - (n+1) \log (n+1)$$

$$= (n+1) \log \frac{n}{n+1} + \frac{A(n) \log n}{n},$$

which tends to -1 as $n \to \infty$, since $\dfrac{A(n) \log n}{n} \to 0$, and

$$(n+1) \log \frac{n}{n+1} = -\log \left[\left(1 + \frac{1}{n}\right)^n \left(1 + \frac{1}{n}\right) \right]$$

tends to -1 as n increases indefinitely.*

Example 1. Consider the series

$$1 + \frac{1 \cdot 3}{2 \cdot 4} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} + \cdot \cdot \cdot \, .$$

Here

$$\frac{u_n}{u_{n+1}} = \frac{2n+2}{2n+1} = 1 + \frac{1}{2n} - \frac{1}{2n(2n+1)}.$$

Since $h = \dfrac{1}{2}$ and $A(n) = \dfrac{-1}{4 + \dfrac{2}{n}}$, the series is divergent.

Example 2. Consider the series

$$\sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot \cdot \cdot (2n-1)}{2 \cdot 4 \cdot \cdot \cdot 2n} \cdot \frac{1}{2n+1}.$$

Now,

$$\frac{u_n}{u_{n+1}} = \frac{(2n+2)(2n+3)}{(2n+1)^2} = \frac{4n^2 + 10n + 6}{4n^2 + 4n + 1}.$$

Dividing the numerator by the denominator gives

$$\frac{u_n}{u_{n+1}} = 1 + \frac{3}{2n} - \frac{1 + \dfrac{3}{2n}}{4n^2 + 4n + 1},$$

---

* See Sec. 10.

and since $h = 3/2$ and

$$A(n) = -\frac{1 + \frac{3}{2n}}{4 + \frac{4}{n} + \frac{1}{n^2}}$$

is bounded when $n \to \infty$, the series is convergent.

PROBLEMS

1. Test for convergence:

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$;

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$;

(c) $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(n-1) + n^2}$;

(d) $\displaystyle\sum_{n=1}^{\infty} \frac{n}{2^n}$.

2. Discuss the convergence of $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots 2n}\right)^p$,

where $p > 2$ and $p \leq 2$.

Ans. Convergent if $p > 2$; divergent if $p \leq 2$.

3. Discuss the convergence of

$$1 + \frac{2 \cdot 4}{1 \cdot 3} + \frac{2 \cdot 4 \cdot 6}{1 \cdot 3 \cdot 5} + \cdots.$$ *Ans.* Divergent.

4. Discuss the convergence of

$$\sum_{n=1}^{\infty} \frac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n-1)} \cdot \frac{1}{2n+2}.$$ *Ans.* Divergent.

5. Discuss the convergence of

$$\sum_{n=1}^{\infty} \left(\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots 2n} \cdot \frac{4n+3}{2n+2}\right)^2.$$

6. Deduce, with the aid of Kummer's test, the following test due to Raabe:

If $\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) > 1$, then the series $\sum_{n=0}^{\infty} u_n$ of positive terms is convergent.

If $\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) < 1$, then the series $\sum_{n=0}^{\infty} u_n$ of positive terms is divergent. This test frequently settles the doubtful case when

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = 1.$$

64. Series of Arbitrary Terms. It follows from Property (C) of Sec. 61 that the theorems of Secs. 62 and 63 are applicable to series all of whose terms are negative. It will be assumed in this section that the terms of the series $\sum_{n=1}^{\infty} a_n$ are arbitrary real numbers, and it will be seen that the behavior of such series is vastly different from that of series whose terms maintain the same sign.

The following theorem, which is basic in all subsequent considerations, is a restatement of the theorem of Sec. 5 on sequences phrased in the language of infinite series.

Theorem. *A necessary and sufficient condition for the convergence of a series* $\sum_{n=1}^{\infty} a_n$ *is that for any* $\epsilon > 0$, *one can find a positive integer* N, *depending on* ϵ, *such that*

$$|s_{n+p} - s_n| \equiv |a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \epsilon,$$

for $n \geq N$ and every integer $p \geq 1$.

It should be carefully noted that the meaning of the inequality

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \epsilon,$$

is that the sum of *any* number of terms of the series $\sum_{n=1}^{\infty} a_n$, beginning with a_{n+1}, must remain less than a preassigned number $\epsilon > 0$, so long as $n \geq N$.

The infinite series

$$a_{n+1} + a_{n+2} + \cdots + a_{n+p} + \cdots = \sum_{k=n+1}^{\infty} a_k,$$

obtained from the given series by discarding the first n terms, is called *the remainder after n terms* and is denoted by the symbol r_n. Thus, if the sum of the series $\sum_{k=1}^{\infty} a_k$ is denoted by S, then

$$s_n + r_n = S, \qquad \text{or} \qquad S - s_n = r_n,$$

and it is clear that r_n represents the error in approximating the sum S of the series by the sum s_n of n terms. The problem of calculating the actual value of the remainder r_n (and, hence, that of the sum S of the series) is, in general, an exceedingly difficult one. However, in many instances it is possible to obtain an estimate of the magnitude of the remainder without determining its actual value. This is illustrated in the following theorem, which is due to Leibnitz.

✓ **Theorem on Alternating Series.** *A series whose terms are alternately positive and negative and such that their absolute values form a monotone null sequence is convergent. The absolute value of the remainder after n terms of such a series does not exceed the absolute value of the* $(n + 1)$*st term.*

In order to prove the theorem, consider the partial sums

$$s_{2n} = a_1 - a_2 + a_3 - \cdots - a_{2n}$$

and

$$s_{2n+2} = s_{2n} + (a_{2n+1} - a_{2n+2}).$$

By hypothesis the terms are monotone decreasing, so that the quantity in the parenthesis is positive. Hence,

$$s_{2n+2} > s_{2n},$$

so that the partial sums of even orders form an increasing sequence. On the other hand,

$$s_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n},$$

and, since the quantity in each parenthesis is positive,

$$s_{2n} < a_1.$$

This states that the increasing sequence $\{s_{2n}\}$ is bounded, and hence, it must converge to some value S. But

$$s_{2n+1} = s_{2n} + a_{2n+1},$$

and

$$\lim_{n\to\infty} s_{2n+1} = \lim_{n\to\infty} s_{2n} + \lim_{n\to\infty} a_{2n+1}.$$

Since the terms of the given series form a null sequence,

$$\lim_{n\to\infty} a_{2n+1} = 0,$$

so that the partial sums of even and odd orders tend to the same limit S.

It remains to show that

$$|r_n| < |a_{n+1}|.$$

Note that if n is even, then

$$\begin{aligned} r_n &= a_{n+1} - (a_{n+2} - a_{n+3}) - (a_{n+4} - a_{n+5}) - \cdots \\ &< a_{n+1}. \end{aligned}$$

If n is odd, then

$$r_n = -a_{n+1} + a_{n+2} - a_{n+3} + \cdots,$$

so that

$$\begin{aligned} -r_n &= a_{n+1} - (a_{n+2} - a_{n+3}) - (a_{n+4} - a_{n+5}) - \cdots \\ &< a_{n+1}. \end{aligned}$$

Thus,

$$|r_n| < |a_{n+1}|,$$

which completes the proof of the theorem.

This theorem enables one to establish easily the convergence of an alternating series, since all that is necessary to show is that

(a) $|a_{n+1}| < |a_n|$, for all $n \geq m$,

(b) $\lim_{n\to\infty} a_n = 0$.

Example 1. The series

$$1 - \frac{1}{2} + \frac{1}{3} - \cdots + (-1)^{n-1}\frac{1}{n} + \cdots$$

is convergent since $|a_{n+1}| < |a_n|$ for all n, and $\lim_{n\to\infty} a_n = 0$. The remainder after 1000 terms, $r_{1000} < 1/1001$, so that the convergence is very slow.

Example 2. The series

$$\frac{5}{2} - \frac{7}{4} + \frac{9}{6} - \frac{11}{8} + \cdots + (-1)^{n-1}\frac{2n+3}{2n} + \cdots$$

has $|a_{n+1}| < |a_n|$ for all n, but it is divergent since $\lim\limits_{n\to\infty} a_n \neq 0$.

Example 3. The series $\frac{1}{2} - \frac{1}{2^2} + \frac{1}{3} - \frac{1}{3^2} + \frac{1}{4} - \frac{1}{4^2} + \cdots$ is alternating, and $\lim\limits_{n\to\infty} a_n = 0$, but it is divergent. Why?

65. Absolute Convergence. *The series* $\sum\limits_{n=1}^{\infty} a_n$ *is said to converge absolutely if* $\sum\limits_{n=1}^{\infty} |a_n|$ *is convergent. The convergent series which do not converge absolutely are called conditionally convergent series.*

The fact that a series may converge in the ordinary sense but fail to converge absolutely can be seen from the following simple example. It was just shown that the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1}\frac{1}{n} + \cdots$$

is convergent, but the series of absolute values is the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} + \cdots,$$

which is known to diverge.

Theorem. *A series* $\sum\limits_{n=1}^{\infty} a_n$ *is convergent if the series of absolute values* $\sum\limits_{n=1}^{\infty} |a_n|$ *is convergent.*

Since the given series is known to converge absolutely, one can find a positive number N, such that

$$|a_{n+1}| + |a_{n+2}| + \cdots + |a_{n+p}| < \epsilon,$$

whenever $n \geq N$, and for every positive integer p. But

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| \leq |a_{n+1}| + |a_{n+2}| + \cdots + |a_{n+p}|;$$

hence,

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \epsilon,$$

which is the criterion for convergence of the series $\sum_{n=1}^{\infty} a_n$.

Inasmuch as the series $\sum_{n=1}^{\infty} |a_n|$ is a series of positive terms, the tests developed in Secs. 62 and 63 can be used to establish the absolute convergence of a series $\sum_{n=1}^{\infty} a_n$. In particular, the tests of Cauchy and d'Alembert can be phrased as follows:

Root Test. *If the terms a_n of the series $\sum_{n=1}^{\infty} a_n$ from some value of n onward satisfy the inequality*

$$\sqrt[n]{|a_n|} \leq r < 1,$$

then the given series converges absolutely.

Corollary. *If* $\lim_{n\to\infty} \sqrt[n]{|a_n|} = r < 1$, *then the series* $\sum_{n\to 1}^{\infty} a_n$ *is absolutely convergent.*

Ratio Test. *If the terms a_n of the series $\sum_{n=1}^{\infty} a_n$, from some value of n onward, satisfy the inequality*

$$\left|\frac{a_{n+1}}{a_n}\right| \leq r < 1,$$

then the given series converges absolutely.

Corollary. *If* $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = r < 1$, *then the series* $\sum_{n=1}^{\infty} a_n$ *is absolutely convergent.*

It was shown above that a series may be convergent, but not absolutely convergent. If, however,

$$\sqrt[n]{|a_n|} \geq 1 \qquad \text{or} \qquad \left|\frac{a_{n+1}}{a_n}\right| \geq 1,$$

then a_n does not tend to zero, so that the tests of Cauchy and d'Alembert can be used also to establish the divergence of alternating series.

As an illustration, let it be required to determine the range of values of x for which the series

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots$$

is convergent.

Now, the test ratio is,

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\dfrac{x^{n+1}}{n+1}}{\dfrac{x^n}{n}}\right| = |x|\frac{n}{n+1}, \sim |x| \frac{1}{1+\frac{1}{n}}$$

and hence,

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = |x|.$$

Therefore, the series converges for all $|x| < 1$ and diverges for all $|x| > 1$. When $x = \pm 1$ the test fails, but substituting $x = 1$ in the series gives

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots,$$

which is convergent. Setting $x = -1$ gives

$$-1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{4} - \cdots$$

which is divergent. Thus, the series converges whenever $-1 < x \leqq 1$.

Series that converge absolutely possess some remarkable properties that are not shared by conditionally convergent series. It will be shown in the next section that the sum of an absolutely convergent series is independent of the order in which the terms of the series are added, but that this is not the case if the series converges conditionally. In fact, by suitably rearranging the terms of a conditionally convergent series, the resulting series can be made to converge to any desired value, or even made to diverge. This striking behavior of conditionally convergent series is illustrated by an example, and the reader will have no difficulty in constructing a general proof along the lines of the example.

Consider the series

$$(65\text{-}1)\quad 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots .$$

The fact that the sum of this series is less than 1 and greater than $\frac{1}{2}$ can be made evident by writing the series as*

$$(1 - \tfrac{1}{2}) + (\tfrac{1}{3} - \tfrac{1}{4}) + (\tfrac{1}{5} - \tfrac{1}{6}) + \cdots ,$$

which shows that $s_n > \frac{1}{2}$ for $n > 2$; on the other hand, the given series can be written as

$$1 - (\tfrac{1}{2} - \tfrac{1}{3}) - (\tfrac{1}{4} - \tfrac{1}{5}) - \cdots ,$$

from which it is clear that $s_n < 1$ for $n \geqq 2$.

Now if the series

$$(65\text{-}2)\quad (1 - \tfrac{1}{2} - \tfrac{1}{4}) + (\tfrac{1}{3} - \tfrac{1}{6} - \tfrac{1}{8}) + (\tfrac{1}{5} - \tfrac{1}{10} - \tfrac{1}{12}) + \cdots$$

is formed by an obvious rearrangement of the terms of (65-1) then the series (65-2) converges to only one-half the sum of the series (65-1). For, adding the first two terms in each parenthesis of (65-2) gives

$$(\tfrac{1}{2} - \tfrac{1}{4}) + (\tfrac{1}{6} - \tfrac{1}{8}) + (\tfrac{1}{10} - \tfrac{1}{12}) + \cdots = \tfrac{1}{2}(1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} - \cdots).$$

It will be shown next that it is possible to rearrange the series

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots ,$$

so as to obtain a new series whose sum is equal to 1. The positive terms of this series in their original order are

$$1,\quad \tfrac{1}{3},\quad \tfrac{1}{5},\quad \tfrac{1}{7},\quad \tfrac{1}{9}, \cdots .$$

The negative terms are

$$-\tfrac{1}{2},\quad -\tfrac{1}{4},\quad -\tfrac{1}{6},\quad -\tfrac{1}{8}, \cdots .$$

In order to form a series which converges to 1, first pick out, in order, as many positive terms as are needed to make their sum equal to or just greater than 1, then pick out just enough negative terms so that the sum of all terms so far chosen will be just less than 1, then more positive terms until the sum is just greater than 1, etc. Thus the partial sums will be

---

* It will be seen in Sec. 76 that the sum of this series is log. 2.

$$
\begin{aligned}
s_1 &= 1, \\
s_2 &= 1 - \tfrac{1}{2} = \tfrac{1}{2}, \\
s_4 &= 1 - \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{5} = \tfrac{31}{30}, \\
s_5 &= 1 - \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{4} = \tfrac{47}{60}, \\
s_7 &= 1 - \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{4} + \tfrac{1}{7} + \tfrac{1}{9} = \tfrac{1307}{1260}, \\
s_8 &= 1 - \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{4} + \tfrac{1}{7} + \tfrac{1}{9} - \tfrac{1}{6} = \tfrac{1093}{1260}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
$$

It is clear that the series formed by this method will have a sum equal to 1.

66. Properties of Absolutely Convergent Series.

Theorem 1 (Commutative Law of Addition for Series). *The rearrangement of the order of the terms in an absolutely convergent series does not alter the sum of the series.*

The theorem will be established first for a series of positive terms. Let the series

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots \tag{66-1}$$

of positive terms be convergent to the sum S, and denote its nth partial sum by s_n. Suppose that the terms of the series (66-1) are rearranged in any way to give a series

$$b_1 + b_2 + b_3 + \cdots + b_n + \cdots , \tag{66-2}$$

and denote the mth partial sum of (66-2) by s'_m. Inasmuch as every term of the series (66-2) appears somewhere in the series (66-1), and since (66-1) is convergent, it is possible to find a positive integer N such that s_n, for $n > N$, will contain all the terms appearing in s'_m and some others which do not appear in s'_m. Then

$$s'_m < s_n < S.$$

It appears that the partial sums s'_m are bounded, and since the sequence $\{s'_m\}$ is increasing, it must approach some limit, say

$$S' \leq S. \tag{66-3}$$

It remains to demonstrate that S' and S are equal. This can be shown by reversing the point of view and regarding the series

(66-1) as some rearrangement of the series (66-2). It was just shown that (66-2) has a sum S'. Therefore, for a sufficiently large m, s'_m will contain all the terms appearing in the partial sum s_n. Hence,

$$s_n < s'_m < S',$$

and, since $\lim_{n\to\infty} s_n = S$, it follows that

$$S \leq S'. \tag{66-4}$$

A comparison of (66-3) and (66-4) leads to the assertion that the inequality sign cannot hold, so that

$$S' = S.$$

Consider next an arbitrary absolutely convergent series $\sum_{n=1}^{\infty} a_n$, and denote its sum by S. Form the series

$$\sum_{n=1}^{\infty} (|a_n| + a_n) \quad \text{and} \quad \sum_{n=1}^{\infty} (|a_n| - a_n).$$

Each term of these series is positive or zero, and they are convergent since $|a_n| \pm a_n \leq 2|a_n|$. Now the difference of these series is

$$\sum_{n=1}^{\infty} (|a_n| + a_n) - \sum_{n=1}^{\infty} (|a_n| - a_n) = 2\sum_{n=1}^{\infty} a_n = 2S. \tag{66-5}$$

Since both the series in the left-hand member of (66-5) are independent of the arrangement of the terms, their difference $2\sum_{n=1}^{\infty} a_n$ is likewise independent, and this establishes the theorem.

Corollary 1. *If* $\sum_{n=1}^{\infty} a_n$ *is absolutely convergent, then any "subseries" formed from it by deleting any number (finite or infinite) of terms is absolutely convergent.*

This follows directly from the foregoing since the partial sums of the subseries are bounded.

Corollary 2. *If* $\sum_{n=1}^{\infty} a_n$ *is an absolutely convergent series, and if* $M_1, M_2, \cdots, M_n, \cdots$ *is any sequence of quantities whose* *numerical values do not increase indefinitely, then the series*

$$a_1M_1 + a_2M_2 + \cdots + a_nM_n + \cdots$$

is absolutely convergent.

This follows upon the application of Theorem 2, Sec. 62.

Example. Consider the series

$$\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - \cdots .$$

This series is absolutely convergent for all values of x because the series

$$\frac{1}{1^2} - \frac{1}{2^2} + \frac{1}{3^2} - \cdots$$

is absolutely convergent and the numerical value of $\cos nx$ never exceeds unity.

Theorem 2. *If* $\sum_{n=1}^{\infty} u_n$ *and* $\sum_{n=1}^{\infty} v_n$ *are any two absolutely convergent series whose sums are* U *and* V, *then they can be multiplied like finite sums to form the series*

$$\begin{aligned} (66\text{-}6) \quad & u_1v_1 + u_1v_2 + u_2v_1 + u_1v_3 + u_2v_2 + u_3v_1 + \cdots \\ & + u_1v_n + u_2v_{n-1} + \cdots + u_nv_1 + \cdots , \end{aligned}$$

which converges absolutely and whose sum is UV.

It is convenient to display the terms of the product series in the form of a rectangular array of numbers

$$\begin{array}{ccc|c|c}
u_1v_1 & u_1v_2 & u_1v_3 & \cdots\ u_1v_n & \cdots \\
\swarrow & \swarrow & & \swarrow & \\
u_2v_1 & u_2v_2 & u_2v_3 & \cdots\ u_2v_n & \cdots \\
\swarrow & & & \swarrow & \\
u_3v_1 & u_3v_2 & u_3v_3 & \cdots\ u_3v_n & \cdots \\
& \swarrow & & & \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\swarrow & & & & \\
u_nv_1 & u_nv_2 & u_nv_3 & \cdots\ u_nv_n & \cdots \\
\hline
\cdots & \cdots & \cdots & \cdots & \cdots ,
\end{array}$$

from which it is seen that the terms entering into (66-6) appear in the diagonals of the array followed from top to bottom. It will suffice to show that any series formed from this array, in such a fashion that each term $u_i v_j$ of the product appears once and only once in the series, is convergent absolutely. Any other series, and in particular (66-6), can be obtained from it by a rearrangement of terms. It is convenient to consider the series formed from the array by following the boundaries of the squares indicated by dotted lines, namely, the series,

$$\begin{aligned} (66\text{-}7) \quad & u_1v_1 + u_1v_2 + u_2v_2 + u_2v_1 + u_1v_3 + u_2v_3 + u_3v_3 \\ & + u_3v_2 + u_3v_1 + \cdots + u_1v_n + u_2v_n + u_3v_n + \cdots \\ & + u_nv_n + u_nv_{n-1} + \cdots + u_nv_1 + \cdots . \end{aligned}$$

It is assumed first that all terms in the series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ are positive. Let s_m denote the mth partial sum of the series (66-7), and U_n and V_n the nth partial sums of the series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$, respectively. Then

$$(66\text{-}8) \qquad U_nV_n \leq s_m \leq U_{n+1}V_{n+1},$$

so long as $n^2 \leq m \leq (n+1)^2$. But

$$\lim_{n\to\infty} U_nV_n = \lim_{n\to\infty} U_{n+1}V_{n+1} = UV,$$

and it follows from (66-8) that

$$\lim_{n\to\infty} s_m = UV.$$

Next let the series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ be any absolutely convergent series and form the series of absolute values

$$(66\text{-}9) \qquad \sum_{n=1}^{\infty} |u_n| \quad \text{and} \quad \sum_{n=1}^{\infty} |v_n|.$$

The product, of the type (66-6), of these series,

$$|u_1v_1| + |u_1v_2| + |u_2v_1| + |u_1v_3| + |u_2v_2| + |u_3v_1| + \cdots ,$$

is convergent, since each term of the series (66-9) is positive and it was shown that the product of such series is convergent. This establishes the absolute convergence of (66-6).

As above, let s_m be the mth partial sum of the convergent series (66-7), some of whose terms may be negative, and denote the limit of s_m as $m \to \infty$ by S. Since

$$\lim_{n \to \infty} s_m = S$$

is independent of the manner in which $m \to \infty$, let m assume the values

$$1^2, 2^2, 3^2, \cdots, n^2, \cdots.$$

Then,

$$\lim_{n \to \infty} s_{n^2} = S. \tag{66-10}$$

But,

$$s_{n^2} = U_n V_n,$$

and

$$\lim_{n \to \infty} s_{n^2} = \lim_{n \to \infty} U_n V_n = UV. \tag{66-11}$$

A comparison of (66-10) and (66-11) shows that the sum of the product series (66-7) is

$$S = UV.$$

The particular form (66-6) of the product of two series is called the *product of Cauchy*, and it derives its particular importance from a consideration of the products of the power series discussed in Chap. VIII. It may be remarked that the condition enunciated in Theorem 2 is sufficient, but by no means necessary. If the series are conditionally convergent, the product of the two series depends on the particular arrangement of the terms in the product series. However, it was shown by Abel that if the series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ are convergent and have for their sums U and V, then if the Cauchy product (66-6) is convergent, its sum is necessarily UV. Mertens has shown that if one of the two convergent series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ is convergent absolutely, then

Cauchy's product converges to UV. The proofs of these assertions of Abel and Mertens are omitted.*

Example. The geometric series

$$S = \sum_{n=0}^{\infty} r^n \tag{66-12}$$

converges absolutely if $|r| < 1$, and its sum is $S = \dfrac{1}{1-r}$. Hence, the product of the series (66-12) by itself, namely,

$$(1 + r + r^2 + \cdots + r^n + \cdots)(1 + r + r^2 + \cdots + r^n + \cdots) = 1 + 2r + 3r^2 + \cdots + nr^{n-1} + \cdots$$

is convergent to $S^2 = \dfrac{1}{(1-r)^2}$.

PROBLEMS

1. Prove that $\sum_{n=1}^{\infty} r^n \sin n\alpha$ converges absolutely for all values of α so long as $|r| < 1$.

2. Prove that $\sum_{n=1}^{\infty} (-1)^n \dfrac{\cos n\alpha}{n^2}$ is absolutely convergent for all values of α.

3. Show that the series

$$\left(1 + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{4}}\right) + \left(\frac{1}{\sqrt{9}} + \frac{1}{\sqrt{11}} - \frac{1}{\sqrt{6}}\right) + \cdots,$$

formed by rearranging the terms of the convergent series

$$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots,$$

is divergent.

Hint: Note that the general term of the given series,

$$\frac{1}{\sqrt{4n-3}} + \frac{1}{\sqrt{4n-1}} - \frac{1}{\sqrt{2n}},$$

---

* See Knopp, K., Theory and Application of Infinite Series, English ed., p. 320.

is positive and greater than

$$\frac{1}{\sqrt{4n}}+\frac{1}{\sqrt{4n}}-\frac{1}{\sqrt{2n}}=\left(1-\frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{n}}.$$

4. For what range of values of x do the series given below converge absolutely:

(*a*) $x-\dfrac{x^3}{3!}+\dfrac{x^5}{5!}-\cdots;$

(*b*) $\dfrac{x}{1\cdot 3}-\dfrac{x^2}{3\cdot 5}+\dfrac{x^3}{5\cdot 7}-\cdots;$

(*c*) $\dfrac{x}{2}-\dfrac{x^2}{2\cdot 2^2}+\dfrac{x^3}{3\cdot 2^3}-\dfrac{x^4}{4\cdot 2^4}+\cdots;$

(*d*) $10x-10^2x^2+10^3x^3-\cdots;$

(*e*) $\dfrac{\sin x}{1^2}-\dfrac{\sin 3x}{3^2}+\dfrac{\sin 5x}{5^2}-\cdots$?

5. Form the Cauchy products of the series (*a*) and (*b*) and the series (*c*) and (*b*) in Prob. 4.

6. For what values of x is the series

$$\frac{x}{x+2}-\left(\frac{x}{x+2}\right)^2+\left(\frac{x}{x+2}\right)^3-\cdots$$

absolutely convergent?

7. For what values of x is the series

$$\frac{2x}{x+4}+\frac{1}{2}\left(\frac{2x}{x+4}\right)^2+\frac{1}{3}\left(\frac{2x}{x+4}\right)^3+\cdots$$

absolutely convergent? *Ans.* $-\frac{4}{3}<x<4$.

67. Double Series. Consider an infinite array of numbers a_{ij}, $(i, j = 1, 2, 3, \cdots)$, arranged in the form of a table

$$\begin{array}{lllll}
a_{11} & a_{12} & a_{13} & a_{14} & \cdots\cdots\cdots \\
a_{21} & a_{22} & a_{23} & a_{24} & \cdots\cdots\cdots \\
a_{31} & a_{32} & a_{33} & a_{34} & \cdots\cdots\cdots \\
a_{41} & a_{42} & a_{43} & a_{44} & \cdots\cdots\cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots\cdots\cdots \\
a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots\cdots\cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots\cdots\cdots
\end{array}$$

where the first subscript on a_{ij} denotes the row, and the second, the column in which the element a_{ij} appears.

One can form the sum of mn terms of the array,

$$s_{mn} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}.$$

Thus,

$$\begin{aligned} s_{34} = \sum_{i=1}^{3} \sum_{j=1}^{4} a_{ij} &= \sum_{i=1}^{3} (a_{i1} + a_{i2} + a_{i3} + a_{i4}) \\ &= a_{11} + a_{12} + a_{13} + a_{14} \\ &+ a_{21} + a_{22} + a_{23} + a_{24} \\ &+ a_{31} + a_{32} + a_{33} + a_{34}. \end{aligned}$$

It may happen that

$$\lim_{\substack{m \to \infty \\ n \to \infty}} s_{mn} = S,$$

where $m \to \infty$ and $n \to \infty$ independently of one another. In such a case the symbol

$$\sum_{i,j=1}^{\infty} a_{ij}$$

is called a *convergent double series.*

A reference to the preceding section shows that the study of the product of two series $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ can be reduced to that of the study of a double series whose terms are $a_{ij} = u_i v_j$.

68. Series of Functions. Uniform Convergence. The remainder of this chapter is concerned with the study of series whose terms are functions of a real variable x. A series of functions of x

$$(68\text{-}1) \qquad u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

will be denoted by the symbol $\sum_{n=1}^{\infty} u_n(x)$, and it will be assumed that the functions $u_n(x)$ are defined in the interval (a, b).

If x in (68-1) is assigned some definite value x_1, there results a series of constants $\sum_{n=1}^{\infty} u_n(x_1)$, which may converge. If the given series is convergent for every value of x in the given interval (a, b), then it is said to be *convergent in the interval* (a, b). A series $\sum_{n=1}^{\infty} u_n(x)$, convergent in (a, b), defines a function of x in that interval, which will be denoted by the symbol $S(x)$. The sum $s_n(x)$ of the first n terms of the series $\sum_{n=1}^{\infty} u_n(x)$ is called the nth *partial sum* of the series, and the statement that the series converges for a given value of x, say $x = x_1$, means that

$$\lim_{n\to\infty} s_n(x_1) = S(x_1). \tag{68-2}$$

The definition of convergence (68-2) can be stated more explicitly as follows:

Definition. *The series* $\sum_{n=1}^{\infty} u_n(x)$ *is convergent for a given value of* x *if for any* $\epsilon > 0$ *one can find a positive integer* N *such that for all values of* $n \geq N$

$$|S(x) - s_n(x)| < \epsilon.$$

If the sum of the terms of the series $\sum_{n=1}^{\infty} u_n(x)$ beginning with $u_{n+1}(x)$ be denoted by $r_n(x)$, then

$$r_n(x) \equiv S(x) - s_n(x) = u_{n+1}(x) + u_{n+2}(x) + \cdots,$$

and the requirement that

$$|S(x) - s_n(x)| < \epsilon$$

is equivalent to the statement that

$$|r_n(x)| < \epsilon.$$

Thus, the assertion that the series $\sum_{n=1}^{\infty} u_n(x)$ is convergent for a given value of $x = x_1$ means that

$$\lim_{n\to\infty} r_n(x_1) = 0.$$

It should be noted carefully that the definition of convergence just given is concerned with convergence at a given point $x = x_1$ of the interval (a, b), so that the number N appearing in the definition depends not only on the magnitude of ϵ, but also on the choice of the point $x = x_1$ in the interval. To illustrate this important remark, consider the series

$$(68\text{-}3)\quad \frac{x}{x+1} + \frac{x}{(x+1)(2x+1)} + \frac{x}{(2x+1)(3x+1)} + \cdots + \frac{x}{[(n-1)x+1](nx+1)} + \cdots$$

in the interval $0 \leq x \leq 1$.

The nth term of this series can be written as

$$\frac{x}{[(n-1)x+1](nx+1)} = \frac{1}{(n-1)x+1} - \frac{1}{nx+1},$$

so that (68-3) can be written in the more convenient form

$$(68\text{-}4)\quad \left(1 - \frac{1}{x+1}\right) + \left(\frac{1}{x+1} - \frac{1}{2x+1}\right) + \cdots + \left(\frac{1}{(n-1)x+1} - \frac{1}{nx+1}\right) + \cdots.$$

It is clear from (68-4) that the sum $s_n(x)$ of n terms of (68-3) is

$$s_n(x) = 1 - \frac{1}{nx+1} = \frac{nx}{nx+1}.$$

Hence,

$$\lim_{n\to\infty} s_n(x) \equiv S(x) = \lim_{n\to\infty} \frac{nx}{nx+1} = 1, \qquad \text{if} \qquad x \neq 0.$$

Thus, if $x \neq 0$,

$$r_n(x) \equiv S(x) - s_n(x) = 1 - \frac{nx}{nx+1} = \frac{1}{nx+1}.$$

Now let it be required to determine the magnitude of N, corresponding to a prescribed ϵ, such that for all $n \geq N$

$$|r_n(x)| < \epsilon.$$

Since

$$r_n(x) = \frac{1}{nx + 1},$$

the condition $|r_n(x)| < \epsilon$ becomes

$$\frac{1}{nx + 1} < \epsilon.$$

Solving for n gives

$$n > \frac{1}{x}\left(\frac{1}{\epsilon} - 1\right), \tag{68-5}$$

from which it is clear that N depends not only on ϵ, but also on the magnitude of x. Thus, if $\epsilon = 0.01$, (68-5) gives $n > \dfrac{99}{x}$; and for $x = \frac{1}{2}$, N can be chosen to be 199. If $\epsilon = 0.01$ but $x = \frac{1}{4}$, N must be at least $4 \cdot (99) + 1 = 397$. It is clear that if ϵ is fixed and x is allowed to assume values nearer and nearer to zero, then the value of N must be chosen larger and larger.

The series (68-3) converges for all values of x in the interval $0 \leq x \leq 1$, but it is impossible to find a single number N which will satisfy the inequality (68-5) uniformly well, that is, for every value of x in the interval $(0, 1)$. However, there are series of functions of x for which it is possible to find one number N, which does not depend on the particular choice of x in the interval and thus depends on ϵ alone.

An example of such a series is

$$S(x) = \frac{x^2}{1 + x} + \left(\frac{2x^2}{1 + 2x} - \frac{x^2}{1 + x}\right) + \cdots + \left(\frac{nx^2}{1 + nx} - \frac{(n - 1)x^2}{1 + (n - 1)x}\right) + \cdots, \tag{68-6}$$

where x is assumed to lie in the interval $0 \leq x \leq 1$. Obviously

$$s_n(x) = \frac{nx^2}{1 + nx},$$

and

$$S(x) = \lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} \frac{nx^2}{1+nx} = x.$$

Consequently,

$$r_n(x) \equiv S(x) - s_n(x) = x - \frac{nx^2}{1+nx} = \frac{x}{1+nx}$$

The requirement that $r_n(x) < \epsilon$ is, in this case,

$$\frac{x}{1+nx} < \epsilon.$$

Solving this inequality for n gives

$$n > \frac{1}{\epsilon} - \frac{1}{x}. \tag{68-7}$$

The inequality (68-7) shows that N again is a function of both ϵ and x, but there is an important distinction in the character of the inequalities (68-5) and (68-7). If $\epsilon = 0.01$ and $x = \frac{1}{2}$, $n > 100 - 2 = 98$, so that N can be chosen to be equal to 99. If $\epsilon = 0.01$ and $x = \frac{1}{4}$, then $n > 100 - 4 = 96$, and N may be any number greater than 96. If $\epsilon = 0.01$ and $x = \frac{3}{4}$, the least value of N is 99. It is clear that N *remains bounded,* regardless of what value of x is chosen in the interval (0, 1), for $N = 100$ will satisfy the inequality (68-7) with $\epsilon = 0.01$ uniformly well for $0 \leq x \leq 1$.

Thus, it is seen that the reason for the difference in the behavior of the two series is that for the first of them N does not remain bounded (ϵ fixed) when x is allowed to assume various values in the interval, whereas in the second case the choice of $N = \frac{1}{\epsilon}$ will serve uniformly well in the interval (0, 1). This leads to an important definition.

Definition of Uniform Convergence. *The series* $\sum_{n=1}^{\infty} u_n(x)$, *defined in the interval* (a, b), *is uniformly convergent in that interval if, for any* $\epsilon > 0$, *there exists a number* N, *independent of the value of* x *in* (a, b), *such that*

$$|S(x) - s_n(x)| < \epsilon,$$

for all values of $n \geq N$.

It should be noted carefully that the concept of uniform convergence is connected inescapably with the interval of convergence. The reader will profit greatly by proving that the series (68-3), which was shown to be nonuniformly convergent in the interval (0, 1), is uniformly convergent in the interval (½, 1). In the latter interval the function $N(x)$ is bounded [see (68-5)], so that one can specify a single number N depending on ϵ only and not on the particular choice of x in the interval (½, 1).

69. Geometric Interpretation of Uniform Convergence. A geometric interpretation of the uniform and nonuniform convergence of series may help to crystallize these important concepts. The graph of the partial sum $s_n(x)$ of the convergent series $S(x) = \sum_{n=1}^{\infty} u_n(x)$ may be called the nth approximation to the curve $y = S(x)$.

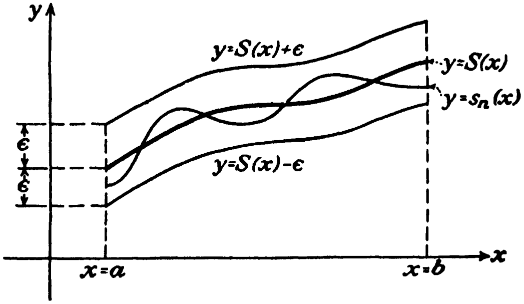

FIG. 67.

If the series is uniformly convergent, then the statement that

$$|r_n(x)| \equiv |S(x) - s_n(x)| < \epsilon, \tag{69-1}$$

whenever $n \geq N$ and for all values of x in (a, b), can be written as follows:

$$S(x) - \epsilon < s_n(x) < S(x) + \epsilon.$$

It follows from these inequalities that the graphs of the approximating functions $y = s_n(x)$ for sufficiently large values of n can be made to lie between the graphs of $y = S(x) + \epsilon$ and $y = S(x) - \epsilon$ for all values of x in (a, b) (Fig. 67).

In the case of a nonuniformly convergent series it is impossible to enclose the approximating curves in a band of width 2ϵ about the graph of $y = S(x)$. To be sure, for any particular value of x, say $x = x_1$, the graph of $y = s_n(x)$ can be made to lie arbitrarily near $y = S(x)$ in the neighborhood of that point, so that the inequality

$$|S(x_1) - s_n(x_1)| < \epsilon$$

is satisfied.

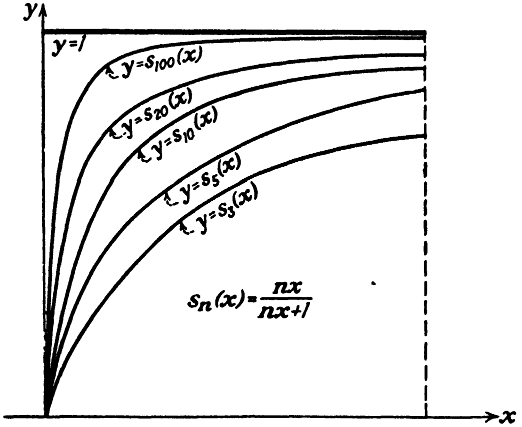

FIG. 68.

In the specific examples of Sec. 68 the series (68–3) is of this second variety. The graph of $S(x)$ is a straight line

$$y = S(x) = 1, \qquad \text{if} \qquad 0 < x \leq 1,$$

and

$$S(0) = 0.$$

The approximating curves

$$y = s_n(x) = \frac{nx}{nx + 1}$$

are shown for several values of n in Fig. 68. All of the approximating curves pass through the origin and approach the line $y = 1$ asymptotically. For any fixed value of x, say $x = x_1$, in the interval (0, 1), it is possible to find a curve $y = s_n(x)$ which will lie as close to $y = S(x_1)$ as desired; but it is obvious that one cannot enclose the approximating curves in a band of width 2ϵ no matter how large a value of n is chosen.

The behavior of the series (68-6) is vastly different (see Fig. 69). The graph of $y = S(x)$ is the straight line $y = x$, and the graphs of

$$y = s_n(x) = \frac{nx^2}{1 + nx}$$

for sufficiently large n can be made to lie entirely within the strip bounded by the lines $y = x \pm \epsilon$.

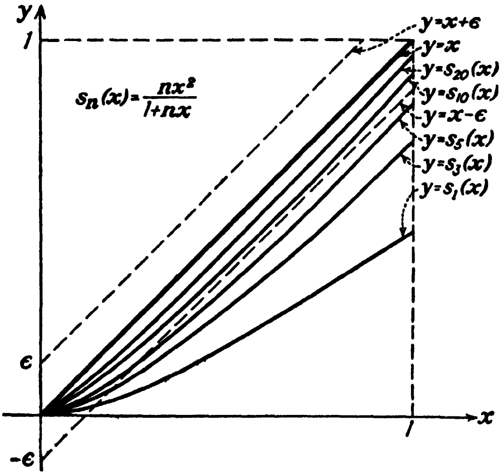

FIG. 69.

The reader will find it instructive to examine carefully the series

$$(69\text{-}2)\quad S(x) = \frac{x}{1 + x^2} + \left(\frac{2^2x}{1 + 2^3x^2} - \frac{x}{1 + x^2}\right) + \left(\frac{3^2x}{1 + 3^3x^2} - \frac{2^2x}{1 + 2^3x^2}\right) + \cdots,$$

for which

$$s_n(x) = \frac{n^2x}{1 + n^3x^2}.$$

Hence, $S(x) = 0$, which is a continuous function for all values of x. The graph of $y = S(x)$ is the entire x-axis, $y = 0$. The graphs of $y = s_n(x)$ for $n = 3, 5, 10$ are indicated in Fig. 70. The approximating curves have peaks in the neighborhood of

$x = 0$ which grow higher with increasing n. Since $y = s_n(x)$ cannot be made to lie within the strip bounded by the lines $y = \pm\epsilon$, it is clear that the convergence cannot be uniform. However, if the study of the series is confined to any interval which does not include $x = 0$, say $0 < \delta \leq x \leq 1$, the series will be found to converge uniformly to the value $S(x) = 0$.

This illustration shows that the sum $S(x)$ of the series may very well be a continuous function even when the series representing the function is not uniformly convergent. However, a discontinuous function cannot be represented by a uniformly convergent series of continuous functions in the neighborhood of the point of discontinuity (see Sec. 70).

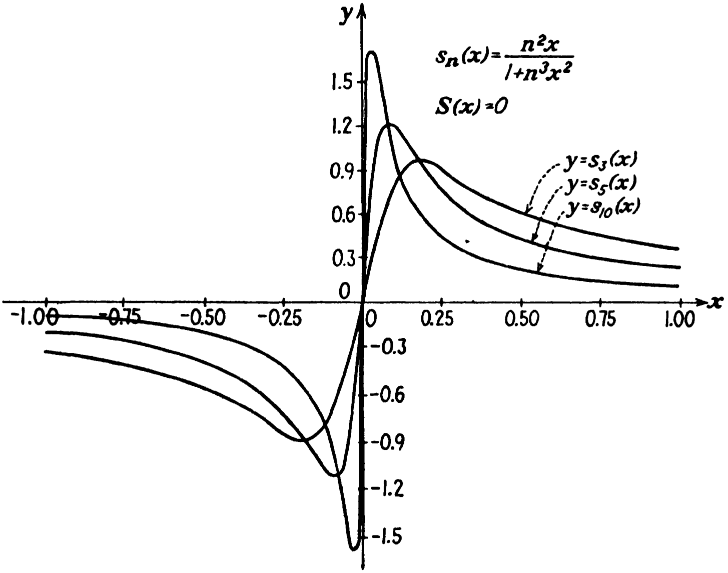

FIG. 70.

PROBLEMS

1. Show that

$$\frac{1}{x+1} - \frac{1}{(x+1)(x+2)} - \cdots - \frac{1}{(x+n-1)(x+n)} - \cdots,$$

for which $s_n(x) = \frac{1}{x+n}$, is uniformly convergent in the interval $0 \leq x \leq 1$

2. Show that the convergence of

$$1 + x + x^2 + x^3 + \cdots + x^n + \cdots$$

is not uniform in the interval $(0, 1)$, but is uniform in the interval $(0, \frac{1}{2})$. Prove that if $|r_n(x)| < \epsilon$, then $n > \dfrac{\log \epsilon + \log (1 - x)}{\log x} - 1$.

3. Show that

$$(1 - x) + x(1 - x) + x^2(1 - x) + \cdots$$

is not uniformly convergent in the interval $0 \leq x \leq 1$.

4 Discuss the nature of the convergence of the series

$$\sum_{n=1}^{\infty} \left(\frac{nx}{1 + n^2x^2} - \frac{(n-1)x}{1 + (n-1)^2x^2} \right),$$

for which $s_n(x) = \dfrac{nx}{1 + n^2x^2}$. Show that $S(x)$ is continuous but that convergence is not uniform in any interval containing the point $x = 0$.

70. Properties of Uniformly Convergent Series. In formulating the definition of uniform convergence, no assumption was made regarding the continuity of the terms of the series. It will be assumed in this section that the functions $u_n(x)$ entering in $\sum_{n=1}^{\infty} u_n(x)$ are continuous functions of x in some interval (a, b). It is clear that the partial sums $s_n(x)$, being the sums of a finite number of continuous functions $u_i(x)$, will be continuous. However, the sum $S(x)$ of an infinite series of continuous functions need not be continuous. This fact is evident from the discussion in the preceding section in connection with the series (68–3). The question naturally arises: Under what circumstances will a series of continuous functions define a continuous function? An answer to this question is given by the following theorem:

Theorem 1. *Let*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

be a series such that each $u_n(x)$ is a continuous function of x in the interval (a, b). If the series is uniformly convergent in (a, b), then the sum of the series is also a continuous function of x in (a, b).

Denote the sum of the series by $S(x)$; then it is required to show that* for any value of x, say $x = x_1$, in the interval (a, b),

$$|S(x_1 + h) - S(x_1)| < \epsilon,$$

whenever $|h| < \delta$. Now

$$|S(x_1 + h) - S(x_1)| = |[S(x_1 + h) - s_n(x_1 + h)] + [s_n(x_1 + h) - s_n(x_1)] + [s_n(x_1) - S(x_1)]|.$$

Since the absolute value of the sum is not greater than the sum of the absolute values, one can write

$$(70\text{-}1) \quad |S(x_1 + h) - S(x_1)| \le |S(x_1 + h) - s_n(x_1 + h)| + |s_n(x_1 + h) - s_n(x_1)| + |s_n(x_1) - S(x_1)|.$$

But, by hypothesis, the given series $\sum_{n=1}^{\infty} u_n(x)$ is uniformly convergent, so that for any prescribed $\epsilon > 0$ one can find a number N such that

$$(70\text{-}2) \quad |s_n(x) - S(x)| < \frac{\epsilon}{3}, \qquad \text{whenever} \qquad n \ge N,$$

and (70-2) must be satisfied for any value of x in the interval (a, b). In particular, it must be true for $x = x_1$ and for $x = x_1 + h$, so that

$$(70\text{-}3) \quad \begin{cases} |s_n(x_1) - S(x_1)| < \dfrac{\epsilon}{3} \\ |s_n(x_1 + h) - S(x_1 + h)| < \dfrac{\epsilon}{3}. \end{cases}$$

Having chosen N so that the inequalities (70-3) are satisfied, choose a number δ so small that for a fixed $n \ge N$

$$(70\text{-}4) \quad |s_n(x_1 + h) - s_n(x_1)| < \frac{\epsilon}{3}, \qquad \text{whenever} \qquad |h| < \delta.$$

This is always possible since $s_n(x)$ is a continuous function. Substituting in the right-hand member of (70-1) gives

$$|S(x_1 + h) - S(x_1)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon, \qquad \text{if} \qquad |h| < \delta.$$

---

* See Sec. 11.

But this is precisely the statement required for the proof of the theorem.

It should be noted that the conditions of Theorem 1 are sufficient to ensure continuity. A nonuniformly convergent series may converge to a continuous function [see the series (69-2)].

Theorem 2. *If a series of continuous functions*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

converges uniformly to $S(x)$ *in* $a \leq x \leq b$, *then*

$$\int_\alpha^\beta S(x)\,dx = \int_\alpha^\beta u_1(x)\,dx + \int_\alpha^\beta u_2(x)\,dx + \cdots + \int_\alpha^\beta u_n(x)\,dx + \cdots,$$

where $a \leq \alpha \leq b$ and $a \leq \beta \leq b$.

This theorem gives a sufficient condition for the integration of a series of functions term by term.

By hypothesis the series $\sum_{n=1}^{\infty} u_n(x)$ is uniformly convergent in the interval (a, b) so that for any $\epsilon > 0$,

$$(70\text{-}5) \qquad |r_n(x)| = |S(x) - s_n(x)| < \epsilon, \qquad \text{for} \qquad n \geq N,$$

and for every x in (a, b). $S(x)$ and $s_n(x)$ denote, as usual, the sum of the series and the nth partial sum. It follows from Theorem 1 that $S(x)$ is continuous, and hence it is integrable in (a, b). Therefore,

$$(70\text{-}6) \qquad \int_\alpha^\beta S(x)\,dx - \int_\alpha^\beta s_n(x)\,dx = \int_\alpha^\beta r_n(x)\,dx,$$

where α and β are any two values of x in (a, b).

From (70-5) it is seen that

$$\int_\alpha^\beta |r_n(x)|\,dx < \epsilon(\beta - \alpha) \leq \epsilon(b - a),$$

and, since

$$\left|\int_\alpha^\beta r_n(x)\,dx\right| \leq \int_\alpha^\beta |r_n(x)|\,dx,$$

(70-6) gives

$$(70\text{-}7) \qquad \left|\int_\alpha^\beta S(x)\,dx - \int_\alpha^\beta s_n(x)\,dx\right| < \epsilon(b - a), \qquad \text{for} \qquad n \geq N.$$

The inequality (70-7) states that the difference between the integral of the sum of the series and the integral of any partial sum $s_n(x)$ can be made arbitrarily small, provided that $n \geq N$. But

$$\int_\alpha^\beta s_n(x)\,dx = \int_\alpha^\beta \sum_{i=1}^{n} u_i(x)\,dx = \sum_{i=1}^{n} \int_\alpha^\beta u_i(x)\,dx,$$

so that (70-7) becomes

$$\left| \int_\alpha^\beta S(x)\,dx - \sum_{i=1}^{n} \int_\alpha^\beta u_i(x)\,dx \right| < \epsilon(b - a), \qquad \text{for} \qquad n \geq N.$$

This inequality, however, states that

$$\int_\alpha^\beta S(x)\,dx = \sum_{i=1}^{\infty} \int_\alpha^\beta u_i(x)\,dx. \tag{70-8}$$

If the limit β in (70-8) is set equal to x, where $a \leq x \leq b$, then the series of integrals

$$\int_a^x u_1(x)\,dx + \int_a^x u_2(x)\,dx + \cdots + \int_a^x u_n(x)\,dx + \cdots$$

converges uniformly to

$$\int_a^x S(x)\,dx.$$

A nonuniformly convergent series, when integrated term by term, may or may not converge to the integral of the function defined by the series. For example, the series (68-3) was shown to be nonuniformly convergent in the interval $0 \leq x \leq 1$. But

$$\int_0^1 S(x)\,dx = \int_0^1 1\,dx = 1.$$

The nth partial sum of (68-3) was found to be

$$s_n(x) = 1 - \frac{1}{nx + 1},$$

and

$$\int_0^1 s_n(x)\,dx = \int_0^1 \left(1 - \frac{1}{nx+1}\right) dx$$
$$= 1 - \frac{\log (n+1)}{n}.$$

Hence,

$$\lim_{n\to\infty} \int_0^1 s_n(x)\,dx = \lim_{n\to\infty} \left(1 - \frac{\log (n+1)}{n}\right) = 1,$$

so that in this case

$$\int_0^1 S(x)\,dx = \sum_{n=1}^{\infty} \int_0^1 u_n(x)\,dx = 1,$$

despite the fact that the given series does not converge uniformly. It is evident from this example that the condition enunciated in Theorem 2 is sufficient, but not necessary, for integration term by term.

As an example of a nonuniformly convergent series that cannot be integrated term by term, consider the series $\sum_{n=1}^{\infty} u_n(x)$ in the interval $0 \le x \le 1$ whose nth partial sum $s_n(x)$ is given by the formula

$$s_n(x) = nxe^{-nx^2}.$$

In this case the sum $S(x)$ of the series is

$$S(x) = \lim_{n\to\infty} nxe^{-nx^2} = 0.$$

Consequently,

$$\int_0^1 S(x)\,dx = 0,$$

whereas

$$\int_0^1 s_n(x)\,dx = \int_0^1 nxe^{-nx^2}\,dx = -\frac{e^{-nx^2}}{2}\bigg|_0^1$$
$$= \frac{1 - e^{-n}}{2}.$$

Thus,

$$\sum_{n=1}^{\infty}\int_0^1 u_n(x)\,dx \equiv \lim_{n\to\infty}\int_0^1 s_n(x)\,dx = \lim_{n\to\infty}\frac{1-e^{-n}}{2} = \frac{1}{2}.$$

The fact that the series $\sum_{n=1}^{\infty} u_n(x)$, whose nth partial sum is $s_n(x) = nxe^{-nx^2}$, cannot be uniformly convergent can be seen

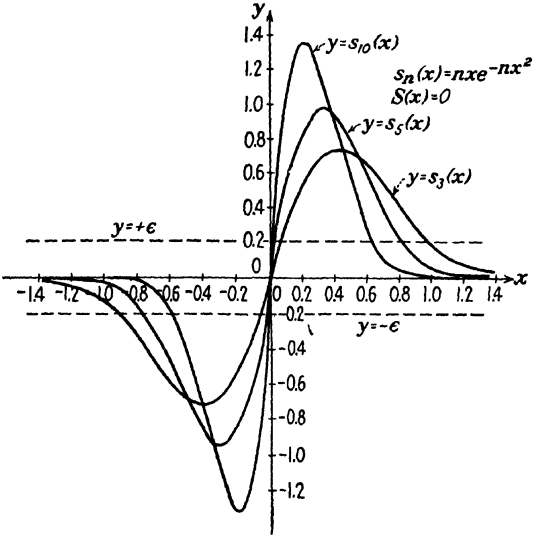

FIG. 71.

from the graphs of the partial sums $y = s_n(x)$ plotted for several values of n in Fig. 71. The peaks of the approximating curves $y = s_n(x)$ grow higher and higher* with the increase in n, and since $S(x) = 0$, it is obvious that it is impossible to draw a band of width 2ϵ about $S(x)$ that will enclose all the approximating curves $y = s_n(x)$ from some value of n onward.

Theorem 3. *If the series*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

---

* The reader will have no difficulty in showing that the maximum value of $s_n(x)$ occurs when $x = \frac{1}{\sqrt{2n}}$, so that $(s_n(x))_{\max.} = \frac{\sqrt{n}}{\sqrt{2e}}$.

is convergent to $S(x)$ *in* (a, b), *and if the derivative of each* $u_n(x)$ *is continuous in* (a, b) *and the series of derivatives*

$$u_1'(x) + u_2'(x) + \cdots + u_n'(x) + \cdots$$

is uniformly convergent in (a, b), *then the series of derivatives converges to* $S'(x)$.

Denote the sum of the series of derivatives by $f(x)$. Then, since the series of derivatives is assumed to be uniformly convergent in (a, b), Theorem 2 permits this series to be integrated term by term so that

$$\int_\alpha^x f(x)\,dx = \int_\alpha^x \sum_{i=1}^{\infty} u_i'(x)\,dx = \sum_{i=1}^{\infty} \int_\alpha^x u_i'(x)\,dx$$

$$= \sum_{i=1}^{\infty} u_i(x)\Big|_\alpha^x = S(x) - S(\alpha),$$

where α and x are in the interval (a, b). Hence,

$$\frac{d}{dx}\int_\alpha^x f(x)\,dx = \frac{d}{dx}[S(x) - S(\alpha)] = S'(x).$$

But $f(x)$ is a continuous function, and therefore,[*]

$$f(x) = S'(x).$$

This theorem provides only sufficient conditions for the differentiation of a series term by term. It is possible to prove the theorem under less severe restrictions. In particular, the requirement that the $u_n'(x)$ be continuous is more severe than it need be.[†]

71. Weierstrass Test for Uniform Convergence. The importance of the concept of uniform convergence is amply illustrated by the theorems of the preceding section, and it is natural to inquire under what circumstances a given series converges uniformly. A direct application of the definition of uniform convergence, leading to the determination of the functional dependence of N upon x and ϵ, is likely to be exceedingly difficult, since it requires knowledge of the expression for the remainder of the series. There are numerous tests, of varying degrees of complexity, but the simplest of these is the test due to K. Weierstrass, which is commonly known as the *Weierstrass M-test.*

---

[*] See Sec. 38.

[†] See Knopp, K., Theory and Application of Infinite Series, p. 342.

Theorem (Weierstrass Test). *Let*

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

be a series of functions such that each $u_i(x)$ is a bounded function of x in (a, b). If there exists a convergent series of positive constants

$$M_1 + M_2 + \cdots + M_n + \cdots$$

such that $|u_i(x)| \leq M_i$ for all values of x in (a, b), then the series

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

is uniformly and absolutely convergent in (a, b).

In order to establish the theorem note that the series of M's is convergent, so that

$$M_{n+1} + M_{n+2} + \cdots + M_{n+p} < \epsilon$$

whenever $n \geq N$, and for all positive integers p. But each $u_i(x)$ is such that $|u_i(x)| \leq M_i$ for all values of x in (a, b); hence,

$$(71\text{-}1) \qquad |u_{n+1}(x)| + |u_{n+2}(x)| + \cdots + |u_{n+p}(x)| < \epsilon.$$

Since (71-1) is independent of the value of x in the interval (a, b), the series $\sum_{i=1}^{\infty} u_i(x)$ is uniformly and absolutely convergent.

It should be noted that the concepts of uniform and absolute convergence are entirely distinct and that they do not imply one another. The series may be uniformly convergent, but not absolutely convergent, and vice versa.

Examples. Consider the series

$$\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \cdots + \frac{\sin nx}{n^2} + \cdots .$$

Since $|\sin nx| \leq 1$, the convergent series

$$\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} + \cdots$$

will serve as an M series for the given series. Therefore, the given series converges uniformly in any interval (a, b), however large. From Theorem 1, Sec. 70, it follows that the series $\sum_{n=1}^{\infty}\frac{\sin nx}{n^2}$ defines a continuous function $S(x)$. Such a series may be integrated term by term in any interval to obtain the integral of $S(x)$. But the series of derivatives

$$\frac{\cos x}{1}+\frac{\cos 2x}{2}+\cdots+\frac{\cos nx}{n}+\cdots$$

is not uniformly convergent. In fact,* it is not even convergent for $x = 2k\pi$, where $k = 0, 1, 2, \cdots$.

On the other hand, the series

$$S_1(x) = \sum_{n=1}^{\infty}\frac{\sin nx}{n^3}$$

upon differentiation term by term gives

$$S_2(x) = \sum_{n=1}^{\infty}\frac{\cos nx}{n^2}.$$

The latter series is obviously uniformly convergent so that the differentiation is legitimate and one is assured that

$$S_2(x) = S_1'(x).$$

72. Abel's Test for Uniform Convergence.

The test of Weierstrass possesses the advantage of great simplicity, but it is applicable only to a restricted class of uniformly convergent series, since every series to which the Weierstrass test is applicable is necessarily absolutely convergent. A more delicate test, essentially due to Abel, is of great practical and theoretical importance, especially in the study of power series.

Theorem (Abel's Test). *The series of functions* $\sum_{n=1}^{\infty}u_n(x)$ *converges uniformly in the interval* $a \leq x \leq b$ *if the functions* $u_n(x)$ *are*

---

* This series, however, converges to $-\log\left(2\sin\frac{x}{2}\right)$ in the interval $0 < x < \pi$. See Prob. 6, Sec. 105.

of the form $u_n(x) = a_n f_n(x)$, *where the constants* a_n *are such that the series* $\sum_{n=1}^{\infty} a_n$ *is convergent and the functions* $f_n(x)$ *are bounded and positive and satisfy the condition*[*]

$$f_{n+1}(x) \leqq f_n(x) < M,$$

for every n *and every value of* x *in* (a, b).

Let

$$\begin{aligned} s_n &= a_1 + a_2 + \cdots + a_n, \\ s'_p &= s_{n+p} - s_n = a_{n+1} + a_{n+2} + \cdots + a_{n+p}, \end{aligned}$$

and

$$S_{n+p} - S_n = u_{n+1} + u_{n+2} + \cdots + u_{n+p} \equiv a_{n+1}f_{n+1} + a_{n+2}f_{n+2} + \cdots + a_{n+p}f_{n+p}.$$

Noting that

$$s'_m - s'_{m-1} = a_{n+m}, \quad \text{if} \quad m > 1, \qquad s'_1 = a_{n+1},$$

one can write

$$\begin{aligned} |S_{n+p} - S_n| &= |a_{n+1}f_{n+1} + a_{n+2}f_{n+2} + \cdots + a_{n+p}f_{n+p}| \\ &= |s'_1 f_{n+1} + (s'_2 - s'_1)f_{n+2} + \cdots + (s'_p - s'_{p-1})f_{n+p}| \\ &= |s'_1(f_{n+1} - f_{n+2}) + s'_2(f_{n+2} - f_{n+3}) + \cdots + s'_p f_{n+p}| \\ &\leqq |s'_1|(f_{n+1} - f_{n+2}) + |s'_2|(f_{n+2} - f_{n+3}) + \cdots + |s'_p| f_{n+p}. \end{aligned}$$

Denote the greatest of the numbers $|s'_m|$, $(m = 1, 2, \cdots, p)$, by k; then

$$(72\text{-}1) \quad |S_{n+p} - S_n| \leq k[(f_{n+1} - f_{n+2}) + (f_{n+2} - f_{n+3}) + \cdots + f_{n+p}] = kf_{n+1}.$$

The series of constants $\sum_{n=1}^{\infty} a_n$ is convergent; hence, for any $\epsilon > 0$ one can find a positive integer N such that for all $n \geq N$ and for every positive integer m,

$$|s_{n+m} - s_n| = |s'_m| < \frac{\epsilon}{M}.$$

---

[*] The proof goes through with obvious minor changes if the sequence of positive functions $f_n(x)$ is monotone increasing with n.

In particular, the inequality is true for the greatest of the s'_m, and it follows that

$$k < \frac{\epsilon}{M}.$$

Substituting this value of k in (72-1) and recalling that $|f_{n+1}| < M$, gives

$$|S_{n+p} - S_n| < \epsilon, \qquad \text{for all} \qquad n \geq N.$$

But N is independent of x, and therefore, $\sum_{n=1}^{\infty} u_n(x)$ is uniformly convergent.

PROBLEMS

1. Test the following series for uniform convergence:

(*a*) $1 + x \cos \alpha + x^2 \cos 2\alpha + x^3 \cos 3\alpha + \cdots, |x| \leq x_1 < 1;$
(*b*) $10x + 10^2x^2 + 10^3x^3 + \cdots;$
(*c*) $\frac{\sin x}{1^2} + \frac{\sin 3x}{3^2} + \frac{\sin 5x}{5^2} + \cdots;$
(*d*) $\frac{1}{2^\alpha} \cos 2x - \frac{1}{3^\alpha} \cos 3x + \frac{1}{4^\alpha} \cos 4x - \cdots;$
(*e*) $\frac{\cos 2x}{1 \cdot 3} + \frac{\cos 4x}{3 \cdot 5} + \frac{\cos 6x}{5 \cdot 7} + \cdots;$
(*f*) $\frac{x}{x-1} + \frac{1}{2}\left(\frac{x}{x-1}\right)^2 + \frac{1}{3}\left(\frac{x}{x-1}\right)^3 + \cdots;$
(*g*) $1 - x + x^2 - x^3 + \cdots;$
(*h*) $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots;$
(*i*) $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots;$
(*j*) $\frac{x-1}{x+1} + \frac{1}{3}\left(\frac{x-1}{x+1}\right)^3 + \frac{1}{5}\left(\frac{x-1}{x+1}\right)^5 + \cdots.$

2. Test the series of derivatives of Prob. 1 for uniform convergence.

CHAPTER VIII

POWER SERIES

73. Power Series. The series of the form

$$\sum_{n=0}^{\infty} a_n x^n \equiv a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots,$$

in which the coefficients of the powers of x are independent of x, occupy an especially prominent place in analysis and are known as *power series*. The totality of values of x for which a power series is convergent is called *the interval of convergence of the series*.

It is readily verified, with the aid of the ratio test, that the series

$$1 + x + x^2 + \cdots + x^n + \cdots$$

converges for all values of x such that $|x| < 1$, and diverges whenever $|x| > 1$, while the series

$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

converges for every real value of x.

Obviously every power series converges for $x = 0$, whatever the coefficients a_n are. The fact that some power series may not converge for any value of x other than zero can be seen from a consideration of the series

$$1 + x + 2!\, x^2 + \cdots + n!\, x^n + \cdots.$$

The ratio of the $(n + 1)$st to the nth term of this series is

$$\frac{u_{n+1}}{u_n} = \frac{n!}{(n-1)!} \frac{x^n}{x^{n-1}} = nx,$$

and since $|nx|$ increases with n for every $x \neq 0$, the series diverges whenever $|x| > 0$.

The following fundamental theorem, due to Abel, furnishes information concerning the character of the convergence of power series.

Theorem. *If a power series* $\sum_{n=0}^{\infty} a_n x^n$ *is convergent for* $x = x_0$, *then it is absolutely convergent for every value of* x *such that* $|x| < |x_0|$. *On the other hand, if the series is divergent for* $x = x_0$, *then it is divergent for every* x *such that* $|x| > |x_0|$.

Consider first a series $\sum_{n=0}^{\infty} a_n x^n$, which converges for some value of x, say $x = x_0$. Then, necessarily, $a_n x_0^n \to 0$ when $n \to \infty$, so that there exists a positive number M such that

$$|a_n x_0^n| < M, \qquad \text{for} \qquad n \geq 0.$$

But

$$|a_n x^n| = \left| a_n x_0^n \cdot \frac{x^n}{x_0^n} \right| = |a_n x_0^n| \cdot \left| \frac{x}{x_0} \right|^n < M \left| \frac{x}{x_0} \right|^n,$$

and, for any x such that $|x| < |x_0|$,

$$\left| \frac{x}{x_0} \right| = r < 1.$$

Accordingly, the series of absolute values,

$$\sum_{n=0}^{\infty} |a_n x^n| \equiv |a_0| + |a_1 x| + |a_2 x^2| + \cdots + |a_n x^n| + \cdots,$$

is term by term less than the convergent geometric series

$$M + Mr + Mr^2 + \cdots + Mr^n + \cdots,$$

and hence, the given series $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely.

The proof of the second part of the theorem follows from the first part. For, assume that the given series diverges for $x = x_0$, and suppose that it converges for some value of x such that $|x| > |x_0|$. Then, by the first part of the theorem, the series must converge for $x = x_0$ as well, which contradicts the hypothesis.

It is possible to establish, with the aid of this theorem, the existence of a number R such that $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for every x such that $|x| < R$ and diverges for all x such that $|x| > R$. However, instead of merely establishing the existence of R, the actual determination of the value of R will be given in the next section. (This determination, of course, establishes the existence of R.) The number R is known as the *radius of convergence* of the series* and the interval $(-R, R)$ is called the *interval of convergence.*

74. Interval of Convergence. The theorem to be established in this section gives the exact determination of the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$. It follows from the application of the root criterion of Cauchy,† that if

$$\sqrt[n]{|a_n|}\,|x| \le r < 1, \tag{74-1}$$

then the series $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for all values of x satisfying the inequality (74-1). In particular, if the limit of $\sqrt[n]{|a_n|}$ exists as $n \to \infty$, and if its magnitude is L, then the series $\sum_{n=0}^{\infty} a_n x^n$ will converge for every x such that

$$L|x| < 1, \tag{74-2}$$

and it will diverge whenever

$$L|x| > 1. \tag{74-3}$$

It follows from (74-2) and (74-3) that the radius of convergence R is given by the formula

$$R = \frac{1}{L} = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}, \tag{74-4}$$

---

* The reason for associating with the number R the term *radius of convergence* is that the region of convergence of a power series where x is a complex number is a circle of radius R.

† See Sec. 62.

provided that $L \neq 0$. But if $L = 0$, the inequality (74-2) is satisfied for all finite values of x, so that in this case the interval of convergence is infinite. If $L = \infty$, it follows from (74-3) that the series diverges for any $x \neq 0$. Thus, if it is agreed, for the purposes of this section, to write $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$, then the formula (74-4) will be applicable to all those series for which the limit L can be determined.

As an illustration of the application of formula (74-4), consider the series

$$(a) \qquad x + x^3 + x^5 + \cdots + x^{2n+1} + \cdots,$$

and

$$(b) \qquad 1 + \left(\frac{x}{2}\right)^2 + \left(\frac{x}{2}\right)^4 + \cdots + \left(\frac{x}{2}\right)^{2n} + \cdots.$$

Since $a_n = 1$ in (a), $\lim_{n\to\infty} \sqrt[n]{a_n} = 1$, and the formula (74-4) gives $R = 1$. In the case of the series (b), $a_{2n} = \frac{1}{2^{2n}}$, so that $\lim_{n\to\infty} \sqrt[2n]{a_{2n}} = \frac{1}{2}$. Consequently, the radius of convergence of (b) is 2.

If one adds the series (a) and (b), there results the series

$$(c) \qquad 1 + x + \left(\frac{x}{2}\right)^2 + x^3 + \left(\frac{x}{2}\right)^4 + \cdots$$

which certainly converges for all values of x such that $|x| < 1$. However, the limit of $\sqrt[n]{|a_n|}$ as $n \to \infty$ does not exist, since $\lim_{n\to\infty} \sqrt[2n]{|a_{2n}|} = \frac{1}{2}$ and $\lim_{n\to\infty} \sqrt[2n+1]{|a_{2n+1}|} = 1$. Formula (74-4) is not applicable in such a case, but it is recalled that every sequence of numbers has a uniquely defined upper limit, and it will be shown next that a generalization of formula (74-4) enables one to calculate the radius of convergence whenever (74-4) is not applicable.

Theorem. *Let L denote the upper limit of the sequence*

$$|a_1|, \sqrt{|a_2|}, \sqrt[3]{|a_3|}, \cdots, \sqrt[n]{|a_n|}, \cdots;$$

then the radius of convergence R of the series $\sum_{n=0}^{\infty} a_n x^n$ *is given by the formula*

$$R = \frac{1}{L} = \frac{1}{\overline{\lim}\, \sqrt[n]{|a_n|}}.$$

Consider first the case where $L = 0$, so that $R = \infty$. If x is any fixed number not equal to zero, then

$$\frac{1}{2|x|} > 0.$$

Since the upper limit L is zero, there are at most a finite number of the $\sqrt[n]{|a_n|}$ that are greater than $L + \epsilon = \dfrac{1}{2|x|}$.* Consequently, one can find a positive integer p such that, for every $n > p$,

$$\sqrt[n]{|a_n|} < \frac{1}{2|x|},$$

or

$$|a_n x^n| < \frac{1}{2^n}.$$

Thus, the absolute values of the terms of the series $\sum_{n=0}^{\infty} a_n x^n$ (for $n > p$) are less than the corresponding terms of the geometric series $\sum_{n=0}^{\infty} \frac{1}{2^n}$ and, since x can be chosen arbitrarily large, it follows that the given series converges for all values of x.

Consider next the case when L is finite but not equal to zero, and let x be any fixed number such that

$$|x| < \frac{1}{L}.$$

Choose any number r between $|x|$ and $\frac{1}{L}$, that is,

$$|x| < r < \frac{1}{L},$$

---

* See Sec. 7.

then, $\frac{1}{r} > L$, so that $\frac{1}{r} = L + \epsilon$. Since only a finite number of the $\sqrt[n]{|a_n|}$ can exceed $L + \epsilon$, it follows that there must exist a positive integer p such that for every $n > p$

$$\sqrt[n]{|a_n|} < \frac{1}{r},$$

or

$$\sqrt[n]{|a_n|} \cdot |x| < \frac{|x|}{r} < 1.$$

Thus,

$$|a_n x^n| < \frac{|x|^n}{r^n} < 1.$$

Again, the general term of the series $\sum_{n=0}^{\infty} a_n x^n$ (for $n > p$) is less in absolute value than the corresponding term of the convergent geometric series $\sum_{n=0}^{\infty} \left|\frac{x}{r}\right|^n$. Consequently the given series must converge absolutely for every value of x satisfying the inequality $|x| < r$. Since r can be chosen as near $\frac{1}{L}$ as desired, the given series will converge for every x such that

$$|x| < \frac{1}{L}.$$

A similar argument shows that if $|x| > \frac{1}{L}$, then $|a_n x^n| > 1$, so that the series cannot converge for any x such that

$$|x| > \frac{1}{L}.$$

In the last possible case, namely, $L = \infty$, the terms of the given series are unbounded if $x \neq 0$. Consequently, the only value of x for which such a series converges is $x = 0$.

In practice it is much easier to determine the radius of convergence with the aid of d'Alembert's test. If the limit as $n \to \infty$ of the ratio $\left|\frac{a_{n+1}}{a_n}\right|$ is defined, then it follows from d'Alembert's test that the series $\sum_{n=0}^{\infty} a_n x^n$ will converge for those values of x for which

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \cdot |x| < 1.$$

Therefore, the given series $\sum_{n=0}^{\infty} a_n x^n$ will converge whenever

$$|x| < \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|},$$

and it will diverge when

$$|x| > \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}.$$

This permits the formulation of the following practical rule for the determination of the interval of convergence:

Rule. *If the series* $\sum_{n=0}^{\infty} a_n x^n$ *is such that*

$$\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = R,$$

then the series converges in the interval $-R < x < R$.

The discussion given above can be extended immediately to cover series in powers of $x - x_0$, namely,

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n. \tag{74-5}$$

The substitution $x' = x - x_0$ in (74-5) gives the power series

$$\sum_{n=0}^{\infty} a_n x'^n,$$

and if the latter series converges for all values of x' such that $|x'| < r$, then the series (74-5) will converge for all values of x

such that $|x - x_0| < r$. Thus, the interval of convergence of the series (74-5) has the point $x = x_0$ as its midpoint.

Examples. The interval of convergence of the series

$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

is infinite since

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} (n+1) = \infty.$$

On the other hand, $\lim\limits_{n\to\infty} \dfrac{a_n}{a_{n+1}}$ does not exist for the series

$$1 + x + \left(\frac{x}{2}\right)^2 + x^3 + \left(\frac{x}{2}\right)^4 + \cdots.$$

However, the sequence

$$|a_1|, \sqrt{|a_2|}, \sqrt[3]{|a_3|}, \ldots, \sqrt[n]{|a_n|}, \ldots$$

is, for this series,

$$1, \tfrac{1}{2}, 1, \tfrac{1}{2}, \ldots,$$

and it is clear that $\overline{\lim}\ \sqrt[n]{|a_n|} = 1$. Hence, the interval of convergence is $-1 < x < 1$.

PROBLEMS

Determine the interval of convergence for the following series, and determine the behavior of the series at the end points of the intervals:

(*a*) $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots;$

(*b*) $10x + 10^2x^2 + 10^3x^3 + \cdots;$

(*c*) $1 + x\cos\theta + x^2\cos 2\theta + \cdots;$

(*d*) $x + 2!x^2 + 3!x^3 + \cdots;$

(*e*) $1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots;$

(*f*) $x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \cdots;$

(*g*) $\frac{x^2}{1\cdot 2} - \frac{x^4}{3\cdot 4} + \frac{x^6}{5\cdot 6} - \cdots;$

(*h*) $1 - 2x + 3x^2 - 4x^3 + \cdots;$

(*i*) $1 + \left(\frac{1}{2}\right)^2 x^2 + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 x^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 x^6 + \cdots;$

(*j*) $\frac{1}{a} + \frac{b}{a^2}x + \frac{b^2}{a^3}x^2 + \frac{b^3}{a^4}x^3 + \cdots.$

75. Properties of Functions Defined by Power Series. The power series $\sum_{n=0}^{\infty} a_n x^n$, whose radius of convergence is R, converges for every value of x such that $|x| < R$ and, hence, defines a function of x in the interior of the interval $(-R, R)$. What can be said about the properties of functions defined by power series? Will such functions be continuous? Is it permissible to integrate and differentiate power series term by term? And if so, what can be said about the intervals of convergence of the resulting series? The answers to these questions furnish an explanation of the prominent role that power series play in analysis.

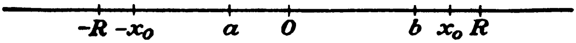

FIG. 72.

Theorem 1. *Let $R > 0$ be the radius of convergence of a power series $\sum_{n=0}^{\infty} a_n x^n$; then this series converges absolutely and uniformly for every value of x in any interval $a \leq x \leq b$ which is interior to the interval $(-R, R)$.*

Inasmuch as the interval (a, b) is assumed to lie entirely within the interval $(-R, R)$, one can choose a positive number x_0 that is greater than $|a|$ and $|b|$ but less than R (Fig. 72). Hence, the interval (a, b) lies entirely within the interval $(-x_0, x_0)$, so that for any value of x in the interval (a, b)

$$|a_n x^n| < |a_n x_0^n|.$$

Accordingly, the series of positive constants

$$\sum_{n=0}^{\infty} |a_n x_0^n|$$

can be used as the Weierstrass M-series to establish the absolute and uniform convergence of the series $\sum_{n=0}^{\infty} a_n x^n$ in the interval (a, b).

Theorem 2. *A power series* $\sum_{n=0}^{\infty} a_n x^n$ *defines a continuous function for all values of* x *in any closed interval* (a, b) *which lies entirely within the interval of convergence of the given series.*

The truth of this statement follows from the preceding theorem upon recalling Theorem 1, Sec. 70.

Theorem 3. *If the radius of convergence of the series* $\sum_{n=0}^{\infty} a_n x^n$ *is* R*, then the radii of convergence of the series* $\sum_{n=0}^{\infty} n a_n x^{n-1}$ *and* $\sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$*, obtained by differentiating and integrating the given series term by term, are also* R.

The proof of the theorem follows from the observation that

$$\overline{\lim}\ \sqrt[n]{|na_n|} = \overline{\lim}\ \sqrt[n]{\left|\frac{a_n}{n+1}\right|} = \overline{\lim}\ \sqrt[n]{|a_n|},$$

since both $\sqrt[n]{n}$ and $\sqrt[n]{\frac{1}{n+1}}$ tend to unity as n increases indefinitely.

Theorem 4. *A power series* $\sum_{n=0}^{\infty} a_n x^n$ *may be differentiated and integrated term by term in any closed interval* (a, b) *which lies entirely within the interval of convergence of the given series.*

The validity of this theorem follows at once from Theorems 1 and 3 of this section, upon noting Theorems 2 and 3, Sec. 70.

76. Abel's Theorem. None of the preceding theorems of this chapter furnishes information concerning the behavior of a power series at the end points of its interval of convergence.

For example, the series

$$(76\text{-}1)\qquad S_1(x) = 1 + x + x^2 + \cdots + x^n + \cdots$$

converges for every value of x interior to the interval $(-1, +1)$ and diverges at $x = \pm 1$.

The series

$$(76\text{-}2) \qquad S_2(x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots$$

has the same interval of convergence as (76-1), but it converges at $x = -1$ and diverges when $x = 1$.

The series obtained by differentiating (76-2) term by term is precisely the series (76-1), and the differentiation is legitimate so long as x lies in the interior of the interval $(-1, 1)$. The series (76-1) defines a continuous function

$$S_1(x) = \frac{1}{1 - x}$$

for every value of x such that $|x| < 1$.

Since

$$\int_0^x \frac{1}{1 - x}\, dx = -\log(1 - x)$$

and

$$\int_0^x (1 + x + x^2 + \cdots + x^n + \cdots)\, dx = x + \frac{x^2}{2} + \cdots + \frac{x^n}{n} + \cdots,$$

one can assert that

$$S_2(x) = -\log(1 - x)$$

so long as $|x| < 1$.

It is important to note that the series (76-1) represents the function

$$S_1(x) = \frac{1}{1 - x}$$

only for those values of x which are interior to the interval $(-1, 1)$, and it is not surprising that the equation

$$\frac{1}{1 - x} = 1 + x + x^2 + \cdots + x^n + \cdots$$

gives a nonsensical result, if one substitutes in it formally $x = -1$. On the other hand, the series (76-2) converges not only for every value of x interior to the interval $(-1, +1)$, but for $x = -1$ as well. In the interior of the interval $(-1, +1)$ the series converges to

$$S_2(x) = -\log(1 - x), \tag{76-3}$$

and the question arises whether formula (76-3) is valid for $x = -1$, so that one is permitted to write

$$S_2(-1) \equiv -1 + \tfrac{1}{2} - \tfrac{1}{3} + \cdots = -\log 2.$$

A theorem, due to Abel, which will be established next, answers this question in the affirmative.

Abel's Theorem. *If a power series* $\sum_{n=0}^{\infty} a_n x^n$, *whose radius of convergence* R *is finite, converges at* $x = R$ *(or* $x = -R$*) then it converges uniformly in the closed interval* $(0, R)$ *(or* $(0, -R)$*).*

There is no loss of generality in assuming that the radius of convergence of the given series is unity. For, a substitution of $x = Ry$ in the series $\sum_{n=0}^{\infty} a_n x^n$, whose radius of convergence is R, produces the series

$$\sum_{n=0}^{\infty} b_n y^n,$$

where $b_n = R^n a_n$, and the series $\sum_{n=0}^{\infty} b_n y^n$ converges in the interval $(-1, 1)$.

Assume that $x = 1$ is an end point of the interval of convergence of the series $\sum_{n=0}^{\infty} a_n x^n$, and that the series of constants

$$a_0 + a_1 + a_2 + \cdots + a_n + \cdots$$

is convergent.

The function $f_n(x) \equiv x^n$ satisfies the conditions of Abel's test (Sec. 72) in the interval $(0, 1)$; hence, the series $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly in the interval $0 \leq x \leq 1$.

The proof of the case when the series converges at $x = -1$ follows from the foregoing if x in the given series is replaced by $-x$.

The theorem of Abel states that if the given power series is known to be convergent at one of the end points of the interval, then the interval of uniform convergence of this series can be extended to include that end point. Since a series of continuous functions that converges uniformly in a given interval defines a continuous function in that interval, it follows that the series $\sum_{n=0}^{\infty} a_n x^n$, which converges at $x = R$, defines a continuous function $S(x)$ in the interval $-R < x \leq R$. Furthermore, since $S(x)$ is continuous at $x = R$, it follows that

$$\lim_{x \to R} S(x) = S(R) = \sum_{n=0}^{\infty} a_n R^n.$$

In the example (76-2) it was found that the series

$$S(x) \equiv x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots$$

converges for $x = -1$. Moreover, in the interval $-1 < x < 1$,

$$S(x) = -\log(1 - x),$$

and the theorem of Abel shows that

$$\lim_{x \to -1} S(x) = \lim_{x \to -1} [-\log(1 - x)] = -\log 2.$$

But

$$\lim_{x \to -1} S(x) = S(-1) = -1 + \tfrac{1}{2} - \tfrac{1}{3} + \cdots.$$

Hence,

$$-\log 2 = -1 + \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{4} - \cdots.$$

77. Uniqueness Theorem on Power Series.

Theorem. *If the power series* $\sum_{n=0}^{\infty} a_n x^n$ *and* $\sum_{n=0}^{\infty} b_n x^n$ *both converge in some interval about the point* $x = 0$ *and have the same sum for every value of* x *in this interval, then the two series are identical.*

By hypothesis

$$(77\text{-}1)\quad a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \cdots = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n + \cdots$$

for all values of x in some interval $(-r, r)$, and since the function $f(x)$ defined by these series is continuous for every point interior to $(-r, r)$, it is certainly continuous at $x = 0$.

Letting $x \rightarrow 0$ in (77-1), there results

$$a_0 = b_0.$$

Dropping a_0 and b_0 in (77-1) and dividing through by x gives, for $x \neq 0$,

$$(77\text{-}2)\quad a_1 + a_2x + a_3x^2 + \cdots = b_1 + b_2x + b_3x^2 + \cdots.$$

Letting $x \rightarrow 0$ in (77-2), there results

$$a_1 = b_1$$

and

$$a_2 + a_3x + a_4x^2 + \cdots = b_2 + b_3x + b_4x^2 + \cdots,$$

for $x \neq 0$.

A continuation of this process leads to the equality

$$a_n = b_n$$

for every value of n.

This theorem asserts that one can equate the corresponding coefficients of powers of x on both sides of the Eq. (77-1). Frequently, this theorem is called the *uniqueness theorem on power series* because it shows that there is not more than one way of representing a given function $f(x)$ in an infinite series of powers of x in a given interval. For, if two different methods lead to two power-series representations of a given function, then the two series are necessarily identical.

78. Algebra of Power Series. It was shown above that a power series $\sum_{n=0}^{\infty} a_nx^n$ converges absolutely for every value of that lies within its interval of convergence. Consequently, it is permissible to rearrange the terms of the power series in any desired manner, so long as x is interior to the interval of convergence.

The following theorems merely restate Property (*D*) of Sec. 61, and Theorem 2, Sec. 66, in a form applicable to power series.

Theorem 1. *If* $\sum_{n=0}^{\infty} a_n x^n$ *and* $\sum_{n=0}^{\infty} b_n x^n$ *are two convergent power series defining, respectively, the functions* $f_1(x)$ *and* $f_2(x)$, *then the series obtained by adding them term by term is the power series*

$$\sum_{n=0}^{\infty} (a_n + b_n)x^n \equiv (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n + \cdots,$$

which converges to $f_1(x) + f_2(x)$ *at least in the common interval of convergence of the two series.*

Example.

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

and

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots .$$

Hence,

$$\frac{1}{1-x} + \cos x = 1 + 1 + x - \frac{x^2}{2!} + x^2 + \frac{x^4}{4!} + \cdots,$$

and the result is, certainly, valid so long as $-1 < x < 1$.

Theorem 2. *The series obtained by multiplying the two convergent series* $\sum_{n=0}^{\infty} a_n x^n$ *and* $\sum_{n=0}^{\infty} b_n x^n$, *defining, respectively, the functions* $f_1(x)$ and $f_2(x)$, *and grouping the terms as for a finite product, is a power series*

$$a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \cdots + (a_0b_n + a_1b_{n-1} + \cdots + a_nb_0)x^n + \cdots,$$

which converges to $f_1(x) \cdot f_2(x)$, *at least in the common interval of convergence of the two original series.*

Example.

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots,$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

and

$$\frac{e^x}{1-x} = 1 + 2x + \frac{5}{2}x^2 + \frac{8}{3}x^3 + \cdots,$$

and the result is valid if $|x| < 1$.

The formulation of a theorem on division of power series is not so simple. Consider the reciprocal of the power series $\sum_{n=0}^{\infty} a_n x^n$, namely,

$$f(x) = \frac{1}{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots}.$$

If $a_0 \neq 0$, one can write

$$\begin{aligned} f(x) &= \frac{1}{a_0\left(1 + \frac{a_1}{a_0}x + \frac{a_2}{a_0}x^2 + \cdots + \frac{a_n}{a_0}x^n + \cdots\right)} \\ &\equiv \frac{1}{a_0}\frac{1}{[1 - (a_1'x + a_2'x^2 + \cdots + a_n'x^n + \cdots)]} \\ &\equiv \frac{1}{a_0}\frac{1}{1-y}, \end{aligned}$$

where $a_n' \equiv -\frac{a_n}{a_0}$ and $y = \sum_{n=1}^{\infty} a_n'x^n$.

For all values of x for which the sum of the series $\sum_{n=1}^{\infty} a_n'x^n$ is numerically less than unity, so that $|y| < 1$,

$$\begin{aligned} (78\text{-}1) \quad f(x) &= \frac{1}{a_0}(1 + y + y^2 + \cdots + y^k + \cdots) \\ &= \frac{1}{a_0}\left[1 + \sum_{n=1}^{\infty} a_n'x^n + \left(\sum_{n=1}^{\infty} a_n'x^n\right)^2 \right. \\ &\qquad \left. + \cdots + \left(\sum_{n=1}^{\infty} a_n'x^n\right)^k + \cdots\right]. \end{aligned}$$

Therefore, the problem is reduced to one that involves multiplication of power series. However, matters are not so simple here, because each term of the series (78-1) is itself an infinite series, and it is impossible to assert that the rearrangement theorem (Sec. 66) established for the case of a single absolutely convergent series is still applicable. If the rearrangement is performed formally, there results a series in powers of x, say $\sum_{n=0}^{\infty} c_n x^n$, and the question arises as to what connection, if any, exists between $f(x)$ and the series so obtained. The reader will guess that the series (78-1) will converge for sufficiently small values of y, so that the expansion $\sum_{n=0}^{\infty} c_n x^n$ will converge to $f(x)$ for small enough values of x. This surmise finds a justification in the following general theorem, the proof of which is omitted:*

Theorem 3. *Given the power series* $f(y) = \sum_{n=0}^{\infty} b_n y^n$ *where*

$$y = \sum_{k=0}^{\infty} a_k x^k;$$

then the power series

$$\sum_{n=0}^{\infty} b_n y^n = \sum_{n=0}^{\infty} b_n \left(\sum_{k=0}^{\infty} a_k x^k \right)^n$$

will certainly converge for every value of x for which the sum of the series $\sum_{k=0}^{\infty} |a_k x^k|$ *is less than the radius of convergence of the series*

$$\sum_{n=0}^{\infty} b_n y^n.$$

An important observation may be made in connection with this theorem. If the series for $f(y)$ in powers of x is to converge, then the radius of convergence R of the series $\sum_{n=0}^{\infty} b_n y^n$ must be such that $R > |a_0|$. If $a_0 = 0$, then $f(y)$ surely can be expanded in a power series in x whenever $R > 0$.

This observation may be clarified by considering an example in which the process of substitution of one infinite series into another is legitimate, and another example in which it is not legitimate.

---

* See KNOPP, K., Theory and Application of Infinite Series, p. 180.

Let it be required to obtain a series in powers of x for the function $e^{\sin x}$ by substituting

$$y = \sin x \equiv x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

into the power series for e^y, namely,

$$e^y \equiv 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots .$$

The intervals of convergence of both of these series are infinite, so that

$$\begin{aligned} e^y = e^{\sin x} &= 1 + \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) \\ &\quad + \frac{1}{2!}\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right)^2 \\ &\quad + \cdots\cdots\cdots\cdots\cdots \\ &= 1 + x + \frac{x^2}{2!} - \frac{3x^4}{4!} - \frac{8x^5}{5!} - \cdots . \end{aligned}$$

This result is valid.

As another illustration, let it be required to obtain a power series in x for the function $\log(1 + e^x)$.

Setting

$$(78\text{-}2) \qquad y = e^x \equiv 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and noting that

$$(78\text{-}3) \qquad \log(1 + y) \equiv y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + \cdots ,$$

gives upon formal substitution of (78-2) in the right-hand member of (78-3), the series

$$(78\text{-}4) \quad \left(1 + x + \frac{x^2}{2!} + \cdots\right) - \frac{1}{2}\left(1 + x + \frac{x^2}{2!} + \cdots\right)^2 + \frac{1}{3}\left(1 + x + \frac{x^2}{2!} + \cdots\right)^3 - \cdots .$$

Denote the sum of the series (78-2) for any particular value of $x > 0$ by r; then (78-4) can be written as

$$r - \tfrac{1}{2}r^2 + \tfrac{1}{3}r^3 - \tfrac{1}{4}r^4 + \cdots,$$

which diverges since r is always greater than unity when $x > 0$. Accordingly, this method will not yield the expansion for

$$f(x) = \log(1 + e^x)$$

in a power series in x.

It follows from Theorem 3 that if the series $y = \sum_{n=0}^{\infty} a_n x^n$ converges for all values of x and the series $\sum_{n=0}^{\infty} b_n y^n$ converges for all values of y, then the substitution of y in terms of x and the rearrangement leading to the series $\sum_{n=0}^{\infty} c_n x^n$ are legitimate.

PROBLEM

Discuss the expansion of $f(x) = \dfrac{1}{1 + e^x}$ in powers of x by using

$$\frac{1}{1+y} = 1 - y + y^2 - y^3 + \cdots,$$

and

$$y = e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

79. Calculations Involving Power Series. It was established in the preceding section that if the radius of convergence R of the series $\sum_{n=0}^{\infty} a_n x^n$ is greater than zero, then one can divide by the power series, provided that $a_0 \neq 0$. Moreover, the resulting expansion will be valid for sufficiently small values of x. Thus

$$\frac{1}{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots} = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n + \cdots. \tag{79-1}$$

A practical scheme for determining the coefficients is outlined next. It follows from (79-1) that

$$1 = (a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots)(c_0 + c_1x + c_2x^2 + \cdots + c_nx^n + \cdots),$$

and an application of the uniqueness theorem on power series gives

$$\begin{aligned}
&a_0c_0 = 1,\\
&a_0c_1 + a_1c_0 = 0,\\
&a_0c_2 + a_1c_1 + a_2c_0 = 0,\\
&a_0c_3 + a_1c_2 + a_2c_1 + a_3c_0 = 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

The first of these equations determines c_0, inasmuch as $a_0 \neq 0$ by hypothesis. A substitution of this value of c_0 in terms of a_0 in the second of these equations yields c_1. Consequently, the coefficients c_i may be determined successively.

As an example, consider the problem of calculating the reciprocal of the series $\sum_{n=1}^{\infty} \frac{x^{n-1}}{n!}$, which has more than illustrative significance. Denote the quotient

$$\frac{1}{1 + \frac{x}{2!} + \frac{x^2}{3!} + \cdots + \frac{x^n}{(n+1)!} + \cdots}$$

by $\sum_{n=0}^{\infty} c_nx^n$, where $c_n \equiv \frac{B_n}{n!}$. Then

$$\left(1 + \frac{x}{2!} + \frac{x^2}{3!} + \cdots + \frac{x^n}{(n+1)!} + \cdots\right)\left(B_0 + \frac{B_1}{1!}x + \frac{B_2}{2!}x^2 + \cdots + \frac{B_n}{n!}x^n + \cdots\right) = 1.$$

and the equations for the determination of the B_n are

$$\begin{aligned}
&B_0 = 1,\\
&\frac{1}{2!}B_0 + \frac{1}{1!}\frac{B_1}{1!} = 0,\\
&\frac{1}{3!}B_0 + \frac{1}{2!}\frac{B_1}{1!} + \frac{1}{1!}\frac{B_2}{2!} = 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots,
\end{aligned}$$

$$\frac{1}{n!}B_0 + \frac{1}{(n-1)!}\frac{B_1}{1!} + \frac{1}{(n-2)!}\frac{B_2}{2!} + \cdots + \frac{1}{1!}\frac{B_{n-1}}{(n-1)!} = 0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

It is readily deduced from these equations that

$$B_1 = -\tfrac{1}{2}, \quad B_2 = \tfrac{1}{6}, \quad B_3 = 0, \quad B_4 = -\tfrac{1}{30},$$
$$B_5 = 0, \quad B_6 = \tfrac{1}{42}, \quad B_7 = 0, \quad B_8 = -\tfrac{1}{30}, \cdots$$

These numbers (or their absolute values) are called *Bernoulli's numbers*, and they play a prominent part in the theory of infinite series.*

It may be remarked that

$$\frac{1}{1 + \frac{x}{2!} + \frac{x^2}{3!} + \cdots + \frac{x^n}{(n+1)!} + \cdots}$$
$$= \frac{x}{\left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots\right) - 1} = \frac{x}{e^x - 1}.$$

Hence,

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n x^n}{n!}.$$

A similar procedure may be employed in calculating the quotient

$$\frac{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots}{b_0 + b_1x + b_2x^2 + \cdots + b_nx^n + \cdots} = c_0 + c_1x + c_2x^2$$
$$+ \cdots + c_nx^n + \cdots,$$

where $b_0 \neq 0$. Clearing of fractions and equating the coefficients of like powers of x gives

$$\begin{aligned}
a_0 &= b_0c_0,\\
a_1 &= b_0c_1 + b_1c_0,\\
a_2 &= b_0c_2 + b_1c_1 + b_2c_0,\\
&\cdots\cdots\cdots\cdots\cdots,\\
a_n &= b_0c_n + b_1c_{n-1} + \cdots + b_nc_0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots,
\end{aligned}$$

---

* See KNOPP, K., Theory and Application of Infinite Series, p. 183.

from which the coefficients c_i (with $i = 0, 1, 2, \cdots$) may be calculated successively. These coefficients are identical with those that would appear if one performed the division of the series by the ordinary rules for division of polynomials arranged in ascending powers of x.

The calculation of the powers of an infinite series $\sum_{n=0}^{\infty} a_n x^n$ may be performed as follows. Let

$$(a_0 + a_1x + \cdots + a_nx^n + \cdots)^m = c_0 + c_1x + \cdots + c_nx^n + \cdots.$$

Then,

$$m \log (a_0 + a_1x + \cdots + a_nx^n + \cdots) = \log (c_0 + c_1x + \cdots + c_nx^n + \cdots).$$

Calculating the derivative and clearing of fractions, one finds,

$$m(a_1 + 2a_2x + \cdots + na_nx^{n-1} + \cdots)(c_0 + c_1x + \cdots + c_nx^n + \cdots) = (a_0 + a_1x + \cdots + a_nx^n + \cdots)(c_1 + 2c_2x + \cdots + nc_nx^{n-1} + \cdots).$$

Equating the coefficients of like powers of x on both sides of this equation gives a system of equations for the determination of $c_1, c_2, \cdots, c_n, \cdots$ in terms of c_0. Clearly $c_0 = a_0^m$, so that the coefficients c_i (where $i = 1, 2, \cdots$) can be determined successively.

PROBLEMS

1. Show by squaring the series that

$$(1 + x + x^2 + \cdots + x^n + \cdots)^2 = 1 + 2x + 3x^2 + \cdots + (n + 1)x^n + \cdots = \frac{1}{(1 - x)^2}.$$

2. If

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$

find the power series for $e^x \sin x$.

3. Show formally that if $f_1(x) = \sum_{n=0}^{\infty} a_n x^n$ and $f_2(x) = \sum_{n=k}^{\infty} b_n x^n$, where $b_k \neq 0$, then the quotient $\dfrac{f_1(x)}{f_2(x)}$ will be of the form

$$\frac{c_{-k}}{x^k} + \frac{c_{-k+1}}{x^{k-1}} + \cdots + \frac{c_{-1}}{x} + c_0 + c_1 x + c_2 x^2 + \cdots .$$

4. *Reversion of Series.* If

$$y - y_0 = a_1(x - x_0) + a_2(x - x_0)^2 + \cdots \tag{1}$$

converges for $|x - x_0| < R$, and if $a_1 \neq 0$, the series may be solved for $x - x_0$ in terms of $y - y_0$ to give

$$x - x_0 = b_1(y - y_0) + b_2(y - y_0)^2 + \cdots . \tag{2}$$

If $\dfrac{y - y_0}{a_1}$ is replaced by z, and $x - x_0$ is set equal to x', the series (1) assumes the form

$$z = a_1'x' + a_2'x'^2 + \cdots + a_n'x'^n + \cdots , \tag{1'}$$

where $a_n' = \dfrac{a_n}{a_1}$, $(n = 1, 2, \cdots)$, and (2) becomes

$$x' = b_1'z + b_2'z^2 + \cdots + b_n'z^n + \cdots , \tag{2'}$$

where

$$b_n' = b_n a_1^n.$$

Substituting (2′) in the right-hand member of (1′) and equating the coefficients of like powers of z, one obtains a set of equations for the determination of the coefficients b_n'.

Carry out the reversion of the series

$$y = \sin x \equiv x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

to obtain

$$x = \sin^{-1} y = y + \frac{1}{2}\frac{y^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{y^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\frac{y^7}{7} + \cdots .$$

5. Use the procedure of Sec. 79 to calculate the first three terms of the series for $\dfrac{1}{e^x + 1}$ to verify the expansion

$$\frac{1}{e^x + 1} = \frac{1}{2} - B_2(2^2 - 1)\frac{x}{2} - B_4(2^4 - 1)\frac{x^3}{4!} - \cdots - B_{2n}(2^{2n} - 1)\frac{x^{2n-1}}{(2n)!} + \cdots .$$

6. Find the first three terms of the series for sec x obtained by calculating

$$\frac{1}{\cos x} = \frac{1}{1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots}.$$

7. Obtain the expansion for

$$\frac{x}{\sin x} = \frac{x}{x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots}.$$

8. Find the series for $\sinh x = \frac{e^x - e^{-x}}{2}$ and $\cosh x = \frac{e^x + e^{-x}}{2}$ with the aid of $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$.

9. Find the expansion for $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ with the aid of

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

10. Assuming

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad |x| < R,$$

show by successive differentiations of $f(x)$ that the coefficients a_n are of the form

$$a_n = \frac{f^{(n)}(0)}{n!},$$

so that

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots.$$

This is the Maclaurin series.

CHAPTER IX

APPLICATIONS OF POWER SERIES

80. Extended Law of the Mean. It was shown in Sec. 20 that a function $f(x)$ that is continuous in the closed interval (a, x) and possesses a derivative at every interior point of the interval can be written in the form

$$f(x) = f(a) + (x - a)f'(\xi),$$

where ξ is some interior point of the interval (a, x). This result, known as the *law of the mean*, can be extended if the function $f(x)$ together with its first $n - 1$ derivatives is continuous in the closed interval (a, b), and if its nth derivative exists at every interior point of (a, b).

In order to obtain the extended law of the mean define the function $F(x)$ by the equation

$$\begin{aligned} F(x) \equiv f(b) - f(x) - (b - x)f'(x) - \frac{(b - x)^2}{2!}f''(x) \\ - \cdots - \frac{(b - x)^{n-1}}{(n - 1)!}f^{(n-1)}(x) - \frac{(b - x)^n}{(b - a)^n}\Big[f(b) - f(a) \\ - (b - a)f'(a) - \frac{(b - a)^2}{2!}f''(a) - \cdots - \frac{(b - a)^{n-1}}{(n - 1)!}f^{(n-1)}(a)\Big], \end{aligned}$$

which is so constructed that

$$F(a) = F(b) = 0.$$

The derivative of $F(x)$ is readily found to be

$$\begin{aligned} \text{(80-1)} \quad F'(x) = -\frac{(b - x)^{n-1}}{(n - 1)!}f^{(n)}(x) \\ + \frac{n(b - x)^{n-1}}{(b - a)^n}\Big[f(b) - f(a) - (b - a)f'(a) - \cdots \\ - \frac{(b - a)^{n-1}}{(n - 1)!}f^{(n-1)}(a)\Big]. \end{aligned}$$

Since the function $F(x)$ satisfies the conditions of Rolle's theorem, $F'(x)$ must vanish for some value of x, say $x = \xi$, interior to the interval (a, b).

Setting $x = \xi$ in (80-1) gives

$$f(b) = f(a) + (b - a)f'(a) + \cdots + \frac{(b - a)^{n-1}}{(n - 1)!}f^{(n-1)}(a) + \frac{(b - a)^n}{n!}f^{(n)}(\xi), \tag{80-2}$$

where $a < \xi < b$. Now, if b in (80-2) is set equal to x, one obtains the formula which expresses the function $f(x)$ as a polynomial of degree n in $x - a$, namely,

$$f(x) = f(a) + (x - a)f'(a) + \cdots + \frac{(x - a)^{n-1}}{(n - 1)!}f^{(n-1)}(a) + \frac{(x - a)^n}{n!}f^{(n)}(\xi), \tag{80-3}$$

where $a < \xi < x$. The coefficients of the powers of $x - a$ in (80-3), except the last one, are constants. The coefficient of $(x - a)^n$ is a function of ξ, which in turn depends upon the magnitude of x. It may happen that this last term in the formula (80-3) is negligibly small in comparison with the preceding ones, and in such a case the function $f(x)$ can be approximately represented by a polynomial of degree $n - 1$ in $x - a$ with constant coefficients.

The formula (80-3) is known as Taylor's formula, and its last term,

$$R_n = \frac{(x - a)^n}{n!}f^{(n)}(\xi), \tag{80-4}$$

is called the *Lagrangian form of the remainder after n terms.*

The foregoing derivation of the Taylor formula is not devoid of some artificiality, and the reader may have a feeling that the definition of the function $F(x)$ anticipates the result. The following section is devoted to a discussion of two simpler methods of deriving formula (80-3), but the gain in simplicity there is achieved by imposing restrictions on the function $f(x)$, which are more severe than those required by the proof given above. It should be noted that the proof just given requires neither the continuity of the nth derivative in (a, b) nor its existence at the end points of the interval.

81. Taylor's Formula. If the function $f(x)$ is assumed to have a continuous nth derivative throughout the closed interval (a, b), it is possible to give a simple derivation of Taylor's formula by integrating the nth derivative of the function n times in succession. Thus, if x is any point of the interval (a, b), then

$$\int_a^x f^{(n)}(x)\,dx = f^{(n-1)}(x)\Big|_a^x = f^{(n-1)}(x) - f^{(n-1)}(a),$$

$$\int_a^x\int_a^x f^{(n)}(x)\,(dx)^2 = \int_a^x f^{(n-1)}(x)\,dx - \int_a^x f^{(n-1)}(a)\,dx$$
$$= f^{(n-2)}(x) - f^{(n-2)}(a) - (x-a)f^{(n-1)}(a),$$

$$\int_a^x\int_a^x\int_a^x f^{(n)}(x)\,(dx)^3 = f^{(n-3)}(x) - f^{(n-3)}(a) - (x-a)f^{(n-2)}(a)$$
$$- \frac{(x-a)^2}{2!}f^{(n-1)}(a),$$

. . . . ,

$$\int_a^x \cdots \int_a^x f^{(n)}(x)\,(dx)^n = f(x) - f(a) - (x-a)f'(a)$$
$$- \frac{(x-a)^2}{2!}f''(a) - \cdots - \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a).$$

Solving for $f(x)$ gives

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n, \tag{81-1}$$

where

$$R_n = \int_a^x \cdots \int_a^x f^{(n)}(x)(dx)^n. \tag{81-2}$$

It is easy to transform the integral form (81-2) of the remainder to the Lagrangian form (80-4) with the aid of the first mean-value theorem for integrals.* In fact,

$$\int_a^x f^{(n)}(x)\,dx = (x-a)f^{(n)}(\xi), \qquad \text{where} \qquad a \leq \xi \leq x.$$

Integrating both members of this equation $n - 1$ times, between the limits a and x, gives

---

* See Sec. 37.

$$R_n = \int_a^x \cdots \int_a^x f^{(n)}(x)\,(dx)^n = \frac{(x-a)^n}{n!} f^{(n)}(\xi),$$

which is the Lagrangian form (80-4) of the remainder.

Another proof of Taylor's formula, which leads to an extremely useful form of the remainder R_n, depends on the integration by parts of the obvious identity

$$(81\text{-}3) \qquad f(a+h) - f(a) = \int_0^h f'(a+h-t)\,dt,$$

where the prime over f denotes the derivative of the function with respect to its argument $a + h - t$.

Integrating (81-3) by parts successively gives

$$\begin{aligned} f(a+h) - f(a) &= \int_0^h f'(a+h-t)\,dt \\ &= tf'(a+h-t)\Big|_0^h + \int_0^h tf''(a+h-t)\,dt \\ &= hf'(a) + \frac{1}{2}t^2 f''(a+h-t)\Big|_0^h \\ &\qquad\qquad + \int_0^h \frac{1}{2}t^2 f'''(a+h-t)\,dt \\ &= \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots . \\ &= hf'(a) + \frac{h^2}{2!}f''(a) + \cdots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a) \\ &\qquad\qquad + \int_0^h \frac{t^{n-1}}{(n-1)!} f^{(n)}(a+h-t)\,dt. \end{aligned}$$

Setting $h = x - a$, one obtains

$$\begin{aligned} f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots \\ + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n, \end{aligned}$$

where

$$(81\text{-}4) \qquad R_n = \frac{1}{(n-1)!}\int_0^{x-a} t^{n-1} f^{(n)}(x-t)\,dt.$$

If the variable of integration in (81-4) be changed to λ, which is defined by the relation $\lambda = x - t$, then (81-4) becomes

$$R_n = \frac{1}{(n-1)!}\int_a^x f^{(n)}(\lambda)(x-\lambda)^{n-1}\,d\lambda. \tag{81-5}$$

In order to obtain the Lagrangian form of the remainder from (81-5), it is only necessary to apply the first mean-value theorem for integrals.* Since $a \leq \lambda \leq x$, $(x-\lambda)^{n-1}$ does not change sign in the interval of integration, and formula (37–3) gives

$$R_n = \frac{f^{(n)}(\xi)}{(n-1)!}\int_a^x (x-\lambda)^{n-1}\,d\lambda = \frac{(x-a)^n}{n!}f^{(n)}(\xi),$$

where $a \leq \xi \leq x$. This establishes the identity of (81-5) with the Lagrangian form (80-4).

A special form of Taylor's formula, which results from setting $a = 0$ in (80-3), is known as the *Maclaurin formula*. It is

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \cdots + \frac{x^{n-1}}{(n-1)!}f^{(n-1)}(0) + R_n, \tag{81-6}$$

where

$$R_n = \frac{f^{(n)}(\xi)}{n!}x^n, \qquad 0 < \xi < x,$$

or

$$R_n = \frac{1}{(n-1)!}\int_a^x f^{(n)}(\lambda)(x-\lambda)^{n-1}\,d\lambda.$$

The Lagrangian form of the remainder is frequently written in a slightly different form. Setting $x - a = h$ and recalling that $a < \xi < x$, one can write

$$\xi = a + \theta h,$$

where θ is some number lying between 0 and 1. The number θ, of course, depends upon the value of x. Consequently, the formula (80-3) can be written as follows:

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(a+\theta h)}{n!}h^n,$$

where $0 < \theta < 1$.

---

* See Sec. 37.

82. Taylor's Series. It was shown in Sec. 80 that Taylor's formula,

$$(82\text{-}1)\quad f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x - a)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x - a)^r,$$

permits one to express a suitably restricted function $f(x)$ as a polynomial of degree n in $x - a$. The first n coefficients of the powers of $x - a$ in (82-1) are constants, whereas the coefficient of $(x - a)^n$ is a function of x, inasmuch as the magnitude of ξ depends upon the choice of x.

It may happen that $f(x)$ is of such a nature that the remainder $R_n = \frac{f^{(n)}(\xi)}{n!}(x - a)^n$ is bounded for all values of x in the interval (a, b), so that

$$|R_n(x)| = \left|\frac{f^{(n)}(\xi)}{n!}(x - a)^n\right| \leq M_n,$$

where M_n is a bound which in general depends on both x and n. If the function $f(x)$ has derivatives of all orders in the interval (a, b), and if R_n tends to zero *uniformly* as to ξ in (a, x) as $n \to \infty$, then the formula (82-1) can be written as the infinite series

$$(82\text{-}2)\quad f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + \cdots$$

which is known as Taylor's series. A special case of Taylor's series (82-2), obtained by setting $a = 0$, is known as Maclaurin's series. It is

$$(82\text{-}3)\quad f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots .$$

The demand that $R_n(x)$ tend to zero uniformly with the increase of n is sufficient to deduce the infinite series representation (82-2) of the function $f(x)$. Stated in precise form this requirement means that for each x in (a, b) and any $\epsilon > 0$, one can find a positive integer N, *independent of the choice of* ξ *in* (a, x), such that $|R_n(x)| < \epsilon$ for all $n \geq N$.

It may happen that a formal application of the formula (82-2) will yield a series which does not represent the function with the aid of which the series is generated. An example, due to Cauchy, may help to clarify this assertion. Let it be required to expand the function

$$\begin{aligned} f(x) &= e^{-\frac{1}{x^2}}, \qquad \text{if} \qquad x \neq 0, \\ &= 0, \qquad\quad \text{if} \qquad x = 0, \end{aligned}$$

in Maclaurin's series. Calculating the derivatives of $f(x)$ gives

$$\begin{aligned} f'(x) &= \frac{2}{x^3}e^{-\frac{1}{x^2}}, \\ f''(x) &= \frac{4 - 6x^2}{x^6}e^{-\frac{1}{x^2}}, \\ &\cdots\cdots\cdots\cdots\cdots, \\ f^{(n)}(x) &= \frac{P(x)}{x^p}e^{-\frac{1}{x^2}}, \end{aligned}$$

where $P(x)$ is a polynomial in x, and $p > 1$.

If the value of $f^{(n)}(x)$ at $x = 0$ is calculated,* it is found that

$$f'(0) = f''(0) = \cdots = f^{(n)}(0) = \cdots = 0.$$

Therefore, the formula (82-3) gives

$$0 + 0 \cdot x + \frac{0}{2!}x^2 + \cdots + \frac{0}{n!}x^n + \cdots,$$

which converges to zero for all values of x, and hence fails to represent the function $e^{-\frac{1}{x^2}}$. Moreover, if $\varphi(x)$ is any function that can be represented by Maclaurin's series as

$$\varphi(x) = \varphi(0) + \varphi'(0)x + \frac{\varphi''(0)}{2!}x^2 + \cdots + \frac{\varphi^{(n)}(0)}{n!}x^n + \cdots,$$

---

* See Sec. 14.

then the function

$$F(x) = \varphi(x) + e^{-\frac{1}{x^2}}$$

will have a Maclaurin expansion that is identical with that for $\varphi(x)$. The reason why the function $f(x) = e^{-\frac{1}{x^2}}$ cannot be represented by a Maclaurin's series is that the remainder,

$$R_n = \frac{P(\xi)}{\xi^p} e^{-\frac{1}{\xi^2}} \frac{x^n}{n!}, \qquad (0 < \xi < x),$$

does not converge to zero for each x in any interval including the point $x = 0$. Of course, Taylor's formula (82-1) can be applied to give a representation of $e^{-\frac{1}{x^2}}$, and in this case the expansion about the origin reduces to a single term, which is precisely the remainder R_n.

It follows from the uniqueness theorem on power series* that if any function can be represented by a power series, then the coefficients in the power series must be identical with those appearing in the Taylor's series representation of that function.

83. Applications of Taylor's Formula.

(*a*) *The Expansion of Sin x.* Let it be required to expand the function $f(x) = \sin x$ in Maclaurin's series. Now,

$$\begin{aligned}
f(x) &= \sin x, & f(0) &= 0;\\
f'(x) &= \cos x, & f'(0) &= 1;\\
f''(x) &= -\sin x, & f''(0) &= 0;\\
f'''(x) &= -\cos x, & f'''(0) &= -1;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\
f^{(n)}(x) &= \sin\left(x + \frac{n\pi}{2}\right), & f^{(n)}(0) &= \sin\frac{n\pi}{2}.
\end{aligned}$$

A substitution in the Maclaurin formula (81-6) gives

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{x^n}{n!}\sin\left(\xi + \frac{n\pi}{2}\right),$$

and, since

$$0 < \xi < x,$$

it follows that

$$\xi = \theta x,$$

---

* See Sec. 77.

where θ is some positive number less than unity. Thus the remainder R_n is

$$R_n = \frac{x^n}{n!}\sin\left(\theta x + \frac{n\pi}{2}\right). \tag{83-1}$$

But the numerical value of sin x never exceeds unity, so that

$$\left|R_n\right| \le \left|\frac{x^n}{n!}\right|$$

for all values of x. Now if x is confined to lie in the interval $(-r, r)$, where r is an arbitrarily large positive number, then

$$\lim_{n\to\infty}\left|\frac{x^n}{n!}\right| = 0$$

for all values of x such that $|x| \le r$. Consequently, $R_n(x)$ converges to zero uniformly in any finite interval $(-r, r)$, and one is justified in writing

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^{n+1}\frac{x^{2n-1}}{(2n-1)!} + \cdots .$$

In order to illustrate the use of the remainder in estimating the magnitude of the error made in calculating the value of sin x with the aid of the infinite series, let it be required to compute the numerical value of sin 10°. Since $10° = \frac{\pi}{18}$ radian, formula (83-1) gives

$$\left|R_n\right| = \left|\frac{1}{n!}\left(\frac{\pi}{18}\right)^n \sin\left(\theta\frac{\pi}{18} + \frac{n\pi}{2}\right)\right| < \frac{1}{n!}\left(\frac{\pi}{18}\right)^n.$$

In particular, if $n = 9$, the polynomial

$$\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}\right)_{x=\frac{\pi}{18}} = \frac{\pi}{18} - \left(\frac{\pi}{18}\right)^3\frac{1}{3!} + \left(\frac{\pi}{18}\right)^5\frac{1}{5!} - \left(\frac{\pi}{18}\right)^7\frac{1}{7!}$$

gives the value of $\sin\frac{\pi}{18}$ with an error that is less than

$$\frac{1}{9!}\left(\frac{\pi}{18}\right)^9.$$

(*b*) *The Expansion of e^x.* A procedure entirely analogous to that employed in the preceding example gives for e^x an expansion

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!}e^{\theta x}, \qquad 0 < \theta < 1.$$

A little reflection shows that for all values of x such that $|x| \leq r$, where r is any positive number,

$$\lim_{n\to\infty} \frac{x^n}{n!} e^{\theta x} = 0.$$

Consequently, one can write

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots. \tag{83-2}$$

For $x = 1$ this formula gives

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \cdots.$$

If the sum of the first n terms is taken as an approximate value of e, the error is less than

$$\frac{1}{n!}\left(1 + \frac{1}{n} + \frac{1}{n^2} + \cdots\right) = \frac{1}{n!}\frac{1}{1 - \frac{1}{n}},$$

since each term of this series is not less than the corresponding term of the series

$$\frac{1}{n!} + \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots.$$

Thus for $n = 10$, the error is less than

$$\frac{1}{10!}\frac{1}{1 - \frac{1}{10}} = 0.0000003,$$

and the value of e is easily found to be 2.718282 correct to six decimal places.

If it is desired to obtain values of e^x when $1 < x < 2$, then the power series in $x - 1$ is more useful than (83-2). The expansion of $f(x)$ in powers of $x - 1$ is

$$f(x) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \cdots + \frac{f^{(n)}(\xi)}{n!}(x-1)^n, \qquad (1 < \xi < x),$$

and, since

$$f^{(n)}(x) = e^x, \qquad f^{(n)}(1) = e, \qquad (n = 0, 1, 2, \cdots),$$

$$\text{(83-3)} \quad e^x = e\left[1 + (x-1) + \frac{1}{2!}(x-1)^2 + \cdots + \frac{1}{(n-1)!}(x-1)^{n-1}\right] + \frac{e^\xi}{n!}(x-1)^n.$$

The remainder in this case is

$$R_n = \frac{e^\xi}{n!}(x-1)^n,$$

so that the error made in using only four terms of the expansion (83-3) is

$$R_4 = \frac{e^\xi}{4!}(x-1)^4, \qquad (1 < \xi < x).$$

If $x = 1.1$, then

$$\begin{aligned} e^{1.1} &= e\left[1 + 0.1 + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!}\right] + \frac{e^\xi}{4!}(0.1)^4 \\ &= 1.1052e + \frac{0.0001}{24}e^\xi. \end{aligned}$$

Since $1 < \xi < 1.1$, and since e^ξ is an increasing function, e^ξ is certainly less than e^2, so that R_4 is certainly less than

$$\frac{0.0001}{24}e^2 < 0.00003.$$

(*c*) *The Binomial Theorem.* The fact that the binomial series

$$\text{(83-4)} \quad 1 + mx + \frac{m(m-1)}{2!}x^2 + \cdots + \frac{m(m-1)\cdots(m-n+1)}{n!}x^n + \cdots$$

converges for all values of x such that $|x| < 1$ is readily established with the aid of the ratio test, but it is not so easy to prove that the series (83-4) actually converges to the value of $(1 + x)^m$ when m is not a positive integer.

This section concludes with the discussion of the convergence of the expansion for the function

$$f(x) = (1 + x)^m.$$

The derivatives of $f(x)$ are found to be

$$\begin{aligned} f'(x) &= m(1 + x)^{m-1}, \\ f''(x) &= m(m - 1)(1 + x)^{m-2}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots, \\ f^{(n)}(x) &= m(m - 1) \cdots (m - n + 1)(1 + x)^{m-n}. \end{aligned}$$

Substituting the values of these derivatives, calculated at $x = 0$, in Maclaurin's formula (81-6) provides the expansion

$$(1 + x)^m = 1 + mx + \frac{m(m - 1)}{2!}x^2 + \cdots + \frac{m(m - 1) \cdots (m - n + 2)}{(n - 1)!}x^{n-1} + R_n,$$

where R_n can be expressed either in the Lagrangian form (80-4) or in the form of the integral (81-5).

It is more convenient to discuss the behavior of R_n by expressing it in the form (81-5), so that

$$\text{(83-5)} \quad R_n = \frac{m(m-1) \cdots (m-n+1)}{(n - 1)!}\int_0^x (1+\lambda)^{m-n}(x-\lambda)^{n-1}\, d\lambda.$$

It follows from the first mean-value theorem for integrals (Sec. 37) that

$$\text{(83-6)} \quad \begin{aligned} \int_0^x (1 + \lambda)^{m-n}(x - \lambda)^{n-1}\, d\lambda &= (1 + \xi)^{m-n}(x - \xi)^{n-1}\int_0^x d\lambda \\ &= (1 + \xi)^{m-n}(x - \xi)^{n-1}x, \end{aligned}$$

where $0 \leq \xi \leq x$.

The substitution of $\xi = \theta x$, where $0 \leq \theta \leq 1$, permits one to express (83-6) in a more convenient form, namely,

$$\begin{aligned} \int_0^x (1 + \lambda)^{m-n}(x - \lambda)^{n-1}\, d\lambda &= (1 + \theta x)^{m-n}(1 - \theta)^{n-1}x^n \\ &= \left(\frac{1 - \theta}{1 + \theta x}\right)^{n-1}(1 + \theta x)^{m-1}x^n, \end{aligned}$$

Hence, (83-5) can be written in the following way:

$$(83\text{-}7)\quad R_n=\left[\frac{(m-1)\cdots(m-n+1)}{(n-1)!}x^{n-1}\right]\cdot\left[\frac{1-\theta}{1+\theta x}\right]^{n-1}\cdot[mx(1+\theta x)^{m-1}]$$

If x is confined to the range $-1 < x < 1$, then the first bracket in (83-7) tends to zero as n becomes infinite.* The second bracket, obviously, does not exceed unity if $|x| < 1$, and the last bracket is independent of n. Accordingly,

$$\lim_{n\to\infty} R_n = 0$$

for all values of x in the closed interval $(-r, r)$, where r is any positive number less than unity. Consequently, it is permissible to write

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \cdots + \frac{m(m-1)\cdots(m-n+1)}{n!}x^n + \cdots .$$

The ratio test shows that this series diverges if the numerical value of x is greater than 1, so that the series does not represent the function $(1+x)^m$ for $|x| > 1$ unless m is a positive integer. If $x = \pm 1$, the ratio test fails, and one must resort to more delicate tests to establish the behavior of the series in this doubtful case.

It follows directly from Kummer's test (Sec. 63), upon setting $a_n = n$, that† the series $\sum_{n=1}^{\infty} u_n$

(*a*) converges absolutely if $\lim_{n\to\infty} n\left(\left|\frac{u_n}{u_{n+1}}\right| - 1\right) > 1$,

and

---

* For, consider the sequence $\{a_n\}$ whose general term a_n is equal to $\frac{(m-1)\cdots(m-n+1)}{(n-1)!}x^{n-1}$. Then $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{m-n}{n}x\right| = |x|$, and it follows from the ratio test that the sequence $\{a_n\}$ is a null sequence if $|x| < 1$.

† This is known as *Raabe's test*. See also Prob. 6, Sec. 63.

$$\text{(b)} \qquad \sum_{n=1}^{\infty} |u_n| \text{ diverges if } \lim_{n\to\infty} n\left(\left|\frac{u_n}{u_{n+1}}\right| - 1\right) < 1.$$

The general term u_n of the binomial series is

$$u_n = \frac{m(m-1)\cdots(m-n+1)}{n!}x^n,$$

so that

$$\frac{u_n}{u_{n+1}} = \frac{n+1}{m-n}\frac{1}{x}.$$

Consider first the case when $x = -1$. Then

$$\frac{u_n}{u_{n+1}} = \frac{n+1}{n-m},$$

so that

$$\lim_{n\to\infty} n\left(\frac{n+1}{n-m} - 1\right) = 1 + m.$$

It follows from (*a*) and (*b*) that the series

converges if $m > 0$,

and

diverges if $m < 0$.

If $x = 1$, the series (83-4) becomes alternating if $n > m + 1$, so that the absolute value of the ratio of the nth term to the $(n + 1)$st is still $\frac{n+1}{n-m}$. The foregoing discussion shows that in this case the series converges absolutely so long as $m > 0$. If $m < 0$, the series of absolute values diverges, but the alternating series may converge conditionally. From the structure of the general term

$$u_n = \frac{m(m-1)\cdots(m-n+1)}{n!}$$

it is clear that if $m \leqq -1$, the successive terms do not decrease, so that the series diverges in this case.

If m is between 0 and -1, one may write the absolute value of u_n in the following form:

$$|u_n| = (1-\alpha)\left(1 - \frac{\alpha}{2}\right)\cdots\left(1 - \frac{\alpha}{n}\right),$$

where $\alpha = 1 + m$. It will be shown next that $\lim_{n\to\infty} |u_n| = 0$, so that the series is convergent (since it is obvious that

$$|u_{n+1}| < |u_n|).$$

Now

$$\log |u_n| = \sum_{k=1}^{n} \log \left(1 - \frac{\alpha}{k}\right) \tag{83-8}$$

and, since $1 - \frac{\alpha}{k}$ is less than unity, each term of the series (83-8) is negative. It is easy to show, with the aid of the integral test of Cauchy, that (83-8) diverges if $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \log |u_n| = -\infty.$$

But if $\log |u_n|$ becomes negatively infinite as $n \to \infty$, then $|u_n|$ itself must tend to zero.

The results just obtained can be summarized as follows: The series (83-4) is convergent if $|x| < 1$ and divergent if $|x| > 1$.

If $x = 1$, convergence is absolute, if $m > 0$.
If $x = 1$, convergence is conditional, if $-1 < m < 0$.
If $x = 1$, the series is divergent, when $m \leq -1$.
If $x = -1$, convergence is absolute if $m > 0$.
If $x = -1$, the series is divergent when $m < 0$.

It follows immediately from Abel's theorem on the continuity of power series (Sec. 76) that if the binomial series converges for $x = 1$, then its sum is precisely equal to 2^m. If $x = -1$ and $m > 0$, then the sum of the series (83-4) is necessarily zero.

PROBLEMS

1. Verify the following expansions:

(a) $\frac{1}{2}(e^x + e^{-x}) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots;$

(b) $\frac{1}{2}(e^x - e^{-x}) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots;$

(c) $e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots;$

(d) $\log x = (x - 1) - \frac{1}{2}(x - 1)^2 + \frac{1}{3}(x - 1)^3 - \cdots;$

(*e*) $\log x = \frac{x-1}{x} + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \cdots$

[*Hint:* Set $x = \frac{1}{z}$ in 1(*d*)];

(*f*) $\tan x = x + \frac{x^3}{3} + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \frac{62}{2835}x^9 + \cdots;$

(*g*) $\sec x = 1 + \frac{x^2}{2!} + \frac{5x^4}{4!} + \frac{61x^6}{6!} + \cdots;$

(*h*) $\sin^{-1}x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \cdots;$

(*i*) $\tan^{-1}x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots;$

(*j*) $\log(x + \sqrt{1+x^2}) = x - \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \cdots;$

(*k*) $e^{\sin x} = 1 + x + \frac{x^2}{2!} - \frac{3x^4}{4!} - \frac{8x^5}{5!} - \frac{3x^6}{6!} + \cdots;$

(*l*) $e^{\sin^{-1}x} = 1 + x + \frac{x^2}{2!} + \frac{2}{3!}x^3 + \frac{5}{4!}x^4 + \cdots.$

2. Expand:

(*a*) $\tan x$ in powers of $x - \frac{\pi}{4}$;

(*b*) e^x in powers of $x - 2$;

(*c*) $\sin x$ in powers of $x - \frac{\pi}{6}$;

(*d*) $2 + x^2 - 3x^5 + 7x^6$ in powers of $x - 1$;

(*e*) $\log x$ in powers of $x - 2$.

84. Euler's Formulas and Hyperbolic Functions. Two particular linear combinations of exponential functions are of such frequent occurrence in mathematics that it has been found convenient to give them a special name. The expression $\frac{1}{2}(e^x + e^{-x})$ is called the *hyperbolic cosine of x*, and is denoted by

$$\cosh x \equiv \tfrac{1}{2}(e^x + e^{-x}).$$

The derivative of $\cosh x$ is equal to $\frac{1}{2}(e^x - e^{-x})$ and is called the *hyperbolic sine of x*. Thus,

$$\sinh x \equiv \tfrac{1}{2}(e^x - e^{-x}).$$

These functions are named hyperbolic because they bear relations to the rectangular hyperbola $x^2 - y^2 = a^2$ that are very similar to those borne by the circular functions to the circle

$x^2 + y^2 = a^2$. The formal analogy between the circular and the hyperbolic functions is best exhibited by the table given later in this section. It will be recalled that the expansion in Maclaurin's series for e^u is

$$e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots, \tag{84-1}$$

so that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{84-2}$$

and

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots. \tag{84-3}$$

Subtracting (84-3) from (84-2) gives

$$e^x - e^{-x} = 2\left(x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right),$$

so that

$$\sinh x \equiv \frac{e^x - e^{-x}}{2} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots. \tag{84-4}$$

On the other hand, addition of (84-2) and (84-3) gives

$$\cosh x \equiv \frac{e^x + e^{-x}}{2} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots. \tag{84-5}$$

Moreover, if it is assumed that (84-1) holds for complex numbers as well as for real numbers, then

$$e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots \tag{84-6}$$

and

$$e^{-ix} = 1 - ix + \frac{(-ix)^2}{2!} + \frac{(-ix)^3}{3!} + \cdots, \tag{84-7}$$

where $i \equiv \sqrt{-1}$. Adding (84-6) and (84-7) and simplifying shows that

$$e^{ix} + e^{-ix} = 2\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\right),$$

which is recognized to be the series for $\cos x$ multiplied by 2. Thus,

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}. \tag{84-8}$$

It is readily verified that subtraction of (84-7) from (84-6) leads to the formula

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}. \tag{84-9}$$

By combining (84-8) with (84-9) there result two interesting relations,

$$\cos x + i \sin x = e^{ix} \qquad \text{and} \qquad \cos x - i \sin x = e^{-ix},$$

which are known as the Euler formulas.

The following table exhibits the formal analogy that exists between the circular and hyperbolic functions; the relations that are given for hyperbolic functions can be established readily from the definitions for the hyperbolic sine and the hyperbolic cosine:

Circular Functions	Hyperbolic Functions
$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$	$\sinh x = \frac{1}{2}(e^{x} - e^{-x})$
$\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$	$\cosh x = \frac{1}{2}(e^{x} + e^{-x})$
$\tan x = \frac{e^{ix} - e^{-ix}}{i(e^{ix} + e^{-ix})}$	$\tanh x = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$
$\cot x = \frac{1}{\tan x}$	$\coth x = \frac{1}{\tanh x}$
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$	$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$	$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$
$\sin^2 x + \cos^2 x = 1$	$\cosh^2 x - \sinh^2 x = 1$
$1 + \tan^2 x = \sec^2 x$	$1 - \tanh^2 x = \operatorname{sech}^2 x$
$\sin 2x = 2 \sin x \cos x$	$\sinh 2x = 2 \sinh x \cosh x$
$\cos 2x = \cos^2 x - \sin^2 x$	$\cosh 2x = \cosh^2 x + \sinh^2 x$
$\sin (x \pm y) = \sin x \cos y \pm \cos x \sin y$	$\sinh (x \pm y) = \sinh x \cosh y \pm \cosh x \sinh y$
$\frac{d \sin x}{dx} = \cos x$	$\frac{d \sinh x}{dx} = \cosh x$
$\frac{d \cos x}{dx} = -\sin x$	$\frac{d \cosh x}{dx} = \sinh x$
$\frac{d \tan x}{dx} = \sec^2 x$	$\frac{d \tanh x}{dx} = \operatorname{sech}^2 x$

85. Integration of Power Series. It was shown in Sec. 75 that a power series converges uniformly for all values of x in any closed interval that is interior to the interval of convergence. Moreover, the series resulting from term-by-term differentiation and integration of a power series have the same intervals of convergence as the original series and will converge to the derivative and integral of the function represented by the given series so long as x is interior to the interval of convergence. These facts are of great use in obtaining the power-series expansions of functions defined by integrals.

Thus, consider the integral

$$\int_0^x \frac{dz}{1+z^2} = \tan^{-1} x.$$

Expanding the integrand in a power series in z, one obtains

$$\begin{aligned}\tan^{-1} x &= \int_0^x (1 - z^2 + z^4 - z^6 + \cdots)\,dz \\ &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots .\end{aligned}$$

Inasmuch as the power series in the integrand is convergent for $|z| < 1$, one is assured that the expansion for $\tan^{-1} x$ is convergent for $|x| < 1$. The reader will convince himself that the procedure employed here in deriving the power series for $\tan^{-1} x$ is much simpler than that of calculating the successive derivatives of $\tan^{-1} x$ and applying the Maclaurin formula. The uniqueness theorem guarantees that the expansion obtained above for $\tan^{-1} x$ is precisely the Maclaurin expansion.

Obviously, the series converges at the end points $x = \pm 1$ of the interval. Since the series converges for $x = 1$, it follows from Abel's theorem (Sec. 76) that the series actually converges to the value $\tan^{-1} 1 = \frac{\pi}{4}$. Thus, one arrives at the conclusion that

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots .$$

Consider another example, namely,

$$\log(1 + x) = \int_0^x \frac{dz}{1+z}$$

$$\log(1+x) = \int_0^x (1 - z + z^2 - \cdots)\,dz$$
$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots.$$

This expansion is valid for $|x| < 1$. For $x = -1$, the series is divergent. For $x = 1$, it converges, and it follows from Abel's theorem that its sum is log 2.

PROBLEMS

1. Show that the remainder in the expansion of $\log(1+x)$ is

$$R_n(x) = (-1)^n \int_0^x \frac{z^n\,dz}{1+z} = \frac{(-1)^n x^{n+1}}{(n+1)(1+\theta x)}.$$

2. Obtain the expansion

$$\sin^{-1} x = \int_0^x \frac{dz}{\sqrt{1-z^2}} = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \cdots,$$

and show that

$$\frac{\pi}{2} = 1 + \frac{1}{2}\cdot\frac{1}{3} + \frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5} + \cdots.$$

3. Show that the remainder in the expansion of $\tan^{-1} x$ in a power series in x is

$$R_n(x) = (-1)^n \int_0^x \frac{z^{2n}}{1+z^2}\,dz = \frac{(-1)^n x^{2n+1}}{(2n+1)(1+(\theta x)^2)},$$

and show that $R_n(1) \to 0$ when $n \to \infty$.

86. Evaluation of Definite Integrals. One of the important uses of infinite series is in the evaluation of some difficult integrals for which the indefinite integral cannot be found in closed form. Three interesting examples of this use of infinite series are given below.

The integral $\int_0^1 \frac{e^x - e^{-x}}{x}\,dx$ cannot be evaluated with the aid of the fundamental theorem since the indefinite integral cannot be obtained in closed form. The expansion for $\frac{e^x - e^{-x}}{x}$, if obtained directly, would lead to an extremely complicated expression for each derivative. The expansion is most easily obtained by using the separate expansions for e^x and e^{-x}. Thus,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots,$$

and

$$e^x - e^{-x} = 2\left(x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right).$$

Hence

$$\int_0^1 \frac{e^x - e^{-x}}{x}\,dx = 2\left(1 + \frac{1}{3 \cdot 3!} + \frac{1}{5 \cdot 5!} + \cdots\right) = 2.1145.$$

In order to evaluate the integral $\int_0^\pi e^{\sin x}\,dx$, recall that

$$e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots,$$

so that

$$e^{\sin x} = 1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots.$$

Then

$$\int_0^\pi e^{\sin x}\,dx = \int_0^\pi \left(1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots\right) dx$$

$$= 2\int_0^{\frac{\pi}{2}} \left(1 + \sin x + \frac{\sin^2 x}{2!} + \frac{\sin^3 x}{3!} + \cdots\right) dx,$$

which can be evaluated with the aid of the Wallis formula,

$$\int_0^{\frac{\pi}{2}} \sin^n x\,dx = \int_0^{\frac{\pi}{2}} \cos^n x\,dx = \frac{(n-1)(n-3)\cdots 2 \text{ or } 1}{n(n-2)\cdots 1 \text{ or } 2}\alpha,$$

where $\alpha = 1$, if n is odd; and $\alpha = \frac{\pi}{2}$, if n is even.

Let it be required to evaluate the integral*

---

* The integrand in this integral becomes discontinuous at $x = 1$, but it is easy to establish the convergence of the integral (see Sec. 95).

$$K = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \qquad \text{where} \qquad k^2 < 1.$$

A substitution of $x = \sin\theta$ reduces the integral to

$$K = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}}.$$

But

$$(1 - k^2 \sin^2\theta)^{-\frac{1}{2}} = 1 + \frac{1}{2}k^2 \sin^2\theta + \frac{1\cdot 3}{2\cdot 4}k^4 \sin^4\theta + \cdots.$$

Consequently,

$$\begin{aligned} K &= \frac{\pi}{2} + \frac{1}{2}k^2 \int_0^{\frac{\pi}{2}} \sin^2\theta\, d\theta + \cdots \\ &\qquad + \frac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots 2n}k^{2n} \int_0^{\frac{\pi}{2}} \sin^{2n}\theta\, d\theta \\ &= \frac{\pi}{2}\left[1 + \left(\frac{1}{2}\right)^2 k^2 + \cdots + \left(\frac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots 2n}\right)^2 k^{2n} + \cdots\right. \end{aligned}$$

since

$$\int_0^{\frac{\pi}{2}} \sin^{2n}\theta\, d\theta = \frac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots 2n}\frac{\pi}{2}.$$

If k is near unity the convergence of the resulting series is not so rapid as may be desired. In such a case a substitution of $l^2 = 1 - k^2$ can be made, so that

$$(1 - k^2 \sin^2\theta)^{-\frac{1}{2}} = \frac{1}{\cos\theta}(1 + l^2 \tan^2\theta)^{-\frac{1}{2}}.$$

Expanding the right-hand member of this expression by the binomial theorem and integrating term by term gives a series that converges rapidly for small values of l.

PROBLEMS

1. Calculate cos 10°, and estimate the maximum error committed by neglecting terms after x^6.

2. Find the interval of convergence of the expansion of e^x in power series in x. Determine the number of terms necessary to compute $e^{1.1}$ accurate to four decimal places from this expansion.

3. Compute sin 33°, correct to four decimal places.

4. Develop the power series for $\tan^{-1} x$ in powers of x, and find the interval of convergence.

5. Expand the integrand of $\int_0^x \frac{dx}{1 + x^2}$ in power series in x, and integrate term by term. Compare the results with that of the preceding problem.

6. Compute $\sqrt[5]{35} = 2(1 + \frac{3}{32})^{1/5}$, correct to five decimal places.

7. Develop the power series in x for $\sin^{-1} x$ and hence establish that

$$\frac{\pi}{6} = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3}\left(\frac{1}{2}\right)^3 + \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{1}{5}\left(\frac{1}{2}\right)^5 + \cdots.$$

8. Differentiate term by term the power series in x for $\sin x$ and thus obtain the power series in x for $\cos x$. What is the interval of convergence of the resulting series?

9. Divide the series

$$\sin x \equiv x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

by the series

$$\cos x \equiv 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots,$$

and thus obtain the series for $\tan x$.

10. Differentiate the series for $\sin^{-1} x$ to obtain the expansion in powers of x for $(1 - x^2)^{-1/2}$. Find the interval of convergence. Is convergence absolute? Investigate the behavior of the series at the end points of the interval of convergence.

11. The integrals

$$F(k, \varphi) = \int_0^\varphi \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}$$

and

$$E(k, \varphi) = \int_0^\varphi \sqrt{1 - k^2 \sin^2 \theta}\, d\theta,$$

where $k^2 < 1$, are called the *elliptic integrals of the first and second kinds*, respectively. Discuss the evaluation of these integrals. The numbers

$F\left(k, \frac{\pi}{2}\right)$ and $E\left(k, \frac{\pi}{2}\right)$ are called the *complete elliptic integrals* and are denoted by the letters K and E, respectively. The values of the functions $F(k, \varphi)$ and $E(k, \varphi)$ are tabulated.*

12. The major and minor axes of an elliptical arch are 200 ft. and 50 ft., respectively. Find the length of the arch. Compute the length of the arch between the points where $x = 0$ and $x = 25$. Make an estimate of the remainder.

13. Show that

$$\int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{\sin x}} = \int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{\cos x}} = \sqrt{2}\int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - \frac{1}{2}\sin^2 \varphi}}.$$

Hint: Set $\sqrt{\cos x} = \cos \varphi$.

14. Show that

$$\int_0^{\frac{\pi}{2}} \frac{\sin^2 \theta \, d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \frac{1}{k^2}(K - E).$$

Hint: $\sin^2 \theta = \frac{1}{k^2} - \frac{1}{k^2}(1 - k^2 \sin^2 \theta)$.

15. Show that

$$K \equiv \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \frac{\pi}{2}\left[1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 k^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 k^6 + \cdots\right]$$

if $k^2 < 1$.

16. Find the length of one arch of the sine curve.

17. Find the length of the portion of $y = \sin x$ lying between $x = 1$ and $x = 2$.

18. Show that the length of arc of an ellipse of semiaxes a and b is given by

$$\begin{aligned} s &= 4a\int_0^{\frac{\pi}{2}} \sqrt{1 - e^2 \sin^2 \theta} \, d\theta \\ &= 2\pi a\left(1 - \frac{e^2}{4} - \frac{3}{64}e^4 - \cdots\right), \end{aligned}$$ where e is the eccentricity.

---

* See, for example, Peirce's A Short Table of Integrals.

Estimate the magnitude of the remainder after three terms.

19. Evaluate in series:

(a) $\int_0^1 \sin (x^2)\, dx$; (b) $\int_0^{1/2} \frac{\sin x\, dx}{\sqrt{1 - x^2}}$;

(c) $\int_0^1 \frac{\sin x}{x}\, dx$; (d) $\int_0^x e^{-x^2}\, dx$;

(e) $\int_0^x \cos (x^2)\, dx$; (f) $\int_0^1 (2 - \cos x)^{-1/2}\, dx$

$$= \int_0^1 \left[2 - \left(1 - 2 \sin^2 \frac{x}{2}\right)\right]^{-1/2} dx$$

$$= \int_0^1 \left(1 + 2 \sin^2 \frac{x}{2}\right)^{-1/2} dx;$$

(g) $\int_0^x \frac{\cos x\, dx}{\sqrt{x}}$; (h) $\int_0^x e^{\tan x}\, dx$.

20. Show, by squaring and adding the power series for sin x and cos x, that

$$\sin^2 x + \cos^2 x = 1.$$

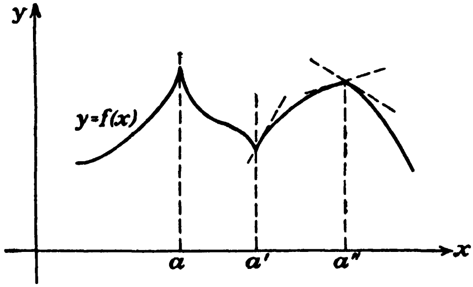

FIG. 73.

87. Maxima and Minima of Functions of One Variable.

A function $f(x)$ is said to have a maximum at $x = a$, if

$$\Delta^+ \equiv f(a + h) - f(a) < 0,$$

and

$$\Delta^- \equiv f(a - h) - f(a) < 0,$$

for all sufficiently small positive values of h. If Δ^+ and Δ^- are both positive for all small positive values of h, then $f(x)$ is said to have a minimum at $x = a$.

It is shown in the elementary calculus that if the function $f(x)$ has a derivative at $x = a$, then the necessary condition for a maximum or a minimum is the vanishing of $f'(x)$ at the point $x = a$. Of course, the function $f(x)$ may attain a maximum or a minimum at $x = a$ without having $f'(a) = 0$, but this can occur only if $f'(x)$ ceases to exist at the critical point (see Fig. 73).

Let it be supposed that $f(x)$ has a continuous derivative of order n in some interval about the point $x = a$. Then it follows from Taylor's formula that

$$\begin{aligned}\Delta^+ &\equiv f(a+h) - f(a)\\ &= f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(a+\theta_1 h)}{n!}h^n,\end{aligned}$$

where $0 < \theta_1 < 1$, and

$$\begin{aligned}\Delta^- &\equiv f(a-h) - f(a)\\ &= -f'(a)h + \frac{f''(a)}{2!}h^2 - \cdots + (-1)^{n-1}\frac{f^{(n-1)}(a)}{(n-1)!}h^{n-1}\\ &\qquad + (-1)^n\frac{f^{(n)}(a-\theta_2 h)}{n!}h^n,\end{aligned}$$

where $0 < \theta_2 < 1$. Let it be assumed further that the first $n - 1$ derivatives of $f(x)$ vanish at $x = a$, but that $f^{(n)}(a)$ is not zero. Then

$$\Delta^+ = \frac{f^{(n)}(a+\theta_1 h)}{n!}h^n$$

and

$$\Delta^- = (-1)^n\frac{f^{(n)}(a-\theta_2 h)}{n!}h^n.$$

Since $f^{(n)}(x)$ is assumed to be continuous in some interval about the point $x = a$, $f^{(n)}(a + \theta_1 h)$ and $f^{(n)}(a - \theta_2 h)$ will have the same sign for sufficiently small values of h.* Consequently, the signs of Δ^+ and Δ^- will be opposite unless n is an even number. But if $f(x)$ is to have a maximum or a minimum at $x = a$, then Δ^+ and Δ^- must be of the same sign. Accordingly, the necessary condition for a maximum or a minimum of $f(x)$

---

* See Theorem 3, Sec. 12.

at $x = a$ is that the first nonvanishing derivative of $f(x)$, at $x = a$, be of even order. Moreover, since both Δ^+ and Δ^- are negative if $f(x)$ is a maximum, it follows that $f^{(n)}(a)$ must be negative. A similar argument shows that, if $f(x)$ has a minimum at $x = a$, then the first nonvanishing derivative of $f(x)$ at $x = a$ must be of even order and positive.

If the first nonvanishing derivative of $f(x)$ at $x = a$ is of odd order, and $f''(a) = 0$, then the point $x = a$ is called a *point of inflection.*

✓ *Example.* Investigate $f(x) = x^5 - 5x^4$ for maxima and minima. Now

$$f'(x) = 5x^4 - 20x^3,$$

which is zero when $x = 0$ and $x = 4$. Then

$$\begin{gathered} f''(x) = 20x^3 - 60x^2, \qquad f''(0) = 0, \qquad f''(4) = 320; \\ f'''(x) = 60x^2 - 120x, \qquad f'''(0) = 0; \\ f^{\text{IV}}(x) = 120x - 120, \qquad f^{\text{IV}}(0) = -120. \end{gathered}$$

Since $f''(4) > 0$, $f(4) = -256$ is a minimum; and since $f^{\text{IV}}(0) < 0$, $f(0) = 0$ is a maximum.

PROBLEMS

1. Examine for maxima and minima:

(*a*) $y = x^4 - 4x^3 + 1$;
(*b*) $y = x^3(x - 5)^2$;
(*c*) $y = x + \cos x$.

2. Find the minimum of the function $y = x^x$, where $x > 0$.

Hint: Consider the minimum of $\log y$. *Ans.* $x = \dfrac{1}{e}$.

3. Show that $x = 0$ gives the minimum value of the function

$$y = e^x + e^{-x} + 2\cos x.$$

88. Taylor's Formula for Functions of Several Variables. Let $f(x, y)$ be a function of two variables x and y that is continuous in the neighborhood of the point (a, b), and that has continuous partial derivatives, up to and including those of order n, in the vicinity of this point.

If a new independent variable t is introduced with the aid of the relations

$$x = a + \alpha t, \qquad y = b + \beta t, \tag{88-1}$$

where α and β are constants, there will result a function of the single variable t, namely,

$$F(t) \equiv f(x, y) = f(a + \alpha t, b + \beta t). \tag{88-2}$$

Expanding $F(t)$ with the aid of the Maclaurin formula gives

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \cdots + \frac{F^{(n)}(\theta t)}{n!}t^n, \tag{88-3}$$

where $0 < \theta < 1$.

It follows from (88-1) and (88-2) that*

$$\begin{aligned} F'(t) &= f_x(x, y)\frac{dx}{dt} + f_y(x, y)\frac{dy}{dt} \\ &= f_x(x, y)\alpha + f_y(x, y)\beta. \end{aligned}$$

Calculating $F''(t)$ and $F'''(t)$ from this expression gives

$$\begin{aligned} F''(t) &= [f_{xx}(x, y)\alpha + f_{yx}(x, y)\beta]\frac{dx}{dt} + [f_{xy}(x, y)\alpha + f_{yy}(x, y)\beta]\frac{dy}{dt} \\ &= f_{xx}(x, y)\alpha^2 + 2f_{xy}(x, y)\alpha\beta + f_{yy}(x, y)\beta^2, \end{aligned}$$

and

$$\begin{aligned} F'''(t) &= [f_{xxx}(x, y)\alpha^2 + 2f_{xyx}(x, y)\alpha\beta + f_{yyx}(x, y)\beta^2]\frac{dx}{dt} \\ &\quad + [f_{xxy}(x, y)\alpha^2 + 2f_{xyy}(x, y)\alpha\beta + f_{yyy}(x, y)\beta^2]\frac{dy}{dt} \\ &= f_{xxx}(x, y)\alpha^3 + 3f_{xxy}(x, y)\alpha^2\beta + 3f_{xyy}(x, y)\alpha\beta^2 + \\ &\qquad f_{yyy}(x, y)\beta^3. \end{aligned}$$

Higher order derivatives of $F(t)$ can be obtained by continuing this process, but the form is evident from those already obtained. Symbolically expressed,

$$\begin{aligned} F'(t) &= \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)f(x, y) \equiv \alpha\frac{\partial f}{\partial x} + \beta\frac{\partial f}{\partial y}, \\ F''(t) &= \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^2 f(x, y) \equiv \alpha^2\frac{\partial^2 f}{\partial x^2} + 2\alpha\beta\frac{\partial^2 f}{\partial x\partial y} + \beta^2\frac{\partial^2 f}{\partial y^2}, \\ F'''(t) &= \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^3 f(x, y) \equiv \alpha^3\frac{\partial^3 f}{\partial x^3} + 3\alpha^2\beta\frac{\partial^3 f}{\partial x^2\partial y} + \\ &\qquad 3\alpha\beta^2\frac{\partial^3 f}{\partial x\partial y^2} + \beta^3\frac{\partial^3 f}{\partial y^3}. \end{aligned}$$

---

* See Sec. 24.

Then

$$F^{(n)}(t) = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^n f(x, y) \equiv \alpha^n\frac{\partial^n f}{\partial x^n} + C_1^n\alpha^{n-1}\beta\frac{\partial^n f}{\partial x^{n-1}\partial y}$$
$$+ \cdots + C_{n-1}^n\alpha\beta^{n-1}\frac{\partial^n f}{\partial x\partial y^{n-1}} + \beta^n\frac{\partial^n f}{\partial y^n},$$

where

$$C_r^n \equiv \frac{n!}{r!(n-r)!}.$$

Since $t = 0$ gives $x = a$ and $y = b$,

$$F(0) = f(a, b), \qquad F'(0) = \alpha f_x(a, b) + \beta f_y(a, b), \cdots.$$

Substituting these expressions in (88-3) gives

$$F(t) \equiv f(x, y) = f(a, b) + [\alpha f_x(a, b) + \beta f_y(a, b)]t$$
$$+ [\alpha^2 f_{xx}(a, b) + 2\alpha\beta f_{xy}(a, b) + \beta^2 f_{yy}(a, b)]\frac{t^2}{2!}$$
$$+ [\alpha^3 f_{xxx}(a,b) + 3\alpha^2\beta f_{xxy}(a,b) + 3\alpha\beta^2 f_{xyy}(a,b) + \beta^3 f_{yyy}(a, b)]\frac{t^3}{3!}$$
$$+ \cdots + R_n,$$

where $R_n = \dfrac{t^n}{n!}\left(\alpha\dfrac{\partial}{\partial x} + \beta\dfrac{\partial}{\partial y}\right)^n f(a + \theta\alpha t, b + \theta\beta t)$.

Since $\alpha t = x - a$ and $\beta t = y - b$, the expansion becomes

$$(88\text{-}4)\quad f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$
$$+ \frac{1}{2!}[f_{xx}(a, b)(x - a)^2 + 2f_{xy}(a, b)(x - a)(y - b)$$
$$+ f_{yy}(a, b)(y - b)^2]$$
$$+ \cdots + R_n.$$

This is Taylor's expansion for a function $f(x, y)$ about the point (a, b). Another useful form of (88-4) is obtained by replacing $x - a$ by h and $y - b$ by k, so that $x = a + h$ and $y = b + k$. Then,

$$(88\text{-}5)\quad f(a + h, b + k) = f(a, b) + f_x(a, b)h + f_y(a, b)k$$
$$+ \frac{1}{2!}[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2]$$
$$+ \cdots + R_n,$$

where $R_n = \dfrac{1}{n!}\left(h\dfrac{\partial}{\partial x} + k\dfrac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k)$.

This formula is frequently written symbolically as

$$\begin{aligned} f(a+h,b+k) = f(a,b) &+ \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a,b) \\ &+ \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a,b) \\ &+ \cdots + R_n. \end{aligned}$$

In particular, if the point (a, b) is $(0, 0)$, the formula (88-4) reads

$$\begin{aligned} \text{(88-6)} \quad f(x,y) = f(0,0) &+ f_x(0,0)x + f_y(0,0)y \\ &+ \frac{1}{2!}[f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y^2] \\ &+ \cdots + R_n, \end{aligned}$$

where $R_n = \dfrac{1}{n!}\left(x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}\right)^n f(\theta x, \theta y), \qquad 0 < \theta < 1.$

This development is known as the *Maclaurin formula for functions of two variables.* It is seen from (88-6) that the Maclaurin formula expresses the function $f(x, y)$ in a series each term of which is a homogeneous polynomial in x and y.

The procedure outlined above can be generalized easily to yield similar expansions for functions of more than two variables.

Example. Obtain the expansion of $\tan^{-1}\dfrac{y}{x}$ about $(1, 1)$ up to the third-degree terms:

$$\begin{aligned} f(x,y) &= \tan^{-1}\frac{y}{x}, & f(1,1) &= \tan^{-1}1 = \frac{\pi}{4}; \\ f_x(x,y) &= -\frac{y}{x^2+y^2}, & f_x(1,1) &= -\frac{1}{2}; \\ f_y(x,y) &= \frac{x}{x^2+y^2}, & f_y(1,1) &= \frac{1}{2}; \\ f_{xx}(x,y) &= \frac{2xy}{(x^2+y^2)^2}, & f_{xx}(1,1) &= \frac{1}{2}; \\ f_{xy}(x,y) &= \frac{y^2-x^2}{(x^2+y^2)^2}, & f_{xy}(1,1) &= 0; \\ f_{yy}(x,y) &= \frac{-2xy}{(x^2+y^2)^2}, & f_{yy}(1,1) &= -\frac{1}{2}. \end{aligned}$$

Then

$$\tan^{-1}\frac{y}{x} = \frac{\pi}{4} - \frac{1}{2}(x-1) + \frac{1}{2}(y-1) + \frac{1}{2!}\left[\frac{1}{2}(x-1)^2 - \frac{1}{2}(y-1)^2\right] + \cdots.$$

PROBLEMS

1. Obtain the expansion for $xy^2 + \cos xy$ about $\left(1, \frac{\pi}{2}\right)$ up to the third-degree terms.

2. Expand $f(x, y) = e^{xy}$ at (1, 1), obtaining three terms.

3. Expand $e^x \cos y$ at (0, 0) up to the fourth-degree terms.

4. Show that for small values of x and y

$$e^x \sin y = y + xy \text{ (approx.)},$$

and

$$e^x \log(1 + y) = y + xy - \frac{y^2}{2} \text{ (approx.)}.$$

5. Expand $f(x, y) = x^3y + x^2y + 1$ about (0, 1).

6. Expand $\sqrt{1 - x^2 - y^2}$ about (0, 0) up to the third-degree terms.

7. Show that the development obtained in Prob. 6 is identical with the binomial expansion of $[1 - (x^2 + y^2)]^{1/2}$.

89. Maxima and Minima of Functions of Several Variables. A function $f(x, y)$ of two independent variables is said to have a maximum at (a, b), if

$$\Delta f \equiv f(a + h, b + k) - f(a, b) < 0$$

for all sufficiently small positive and negative values of h and k. It is said to attain a minimum at (a, b), if

$$\Delta f \equiv f(a + h, b + k) - f(a, b) > 0$$

for all sufficiently small positive and negative values of h and k.

Let it be assumed that $f(x, y)$ attains a maximum (or a minimum) at (a, b). Then the function of a single variable $f(x, b)$ must attain a maximum (or a minimum) at a, and its derivative, if it exists, must vanish at $x = a$. Hence, a necessary condition for a maximum (or a minimum) is that

$$\frac{\partial f}{\partial x} = 0 \quad \text{at} \quad x = a, \tag{89-1}$$

provided that this derivative exists. A similar consideration of the function $f(a, y)$ leads to the conclusion that

$$\frac{\partial f}{\partial y} = 0 \quad \text{at} \quad y = b \tag{89-2}$$

whenever the derivative exists.

The equations (89-1) and (89-2) may be solved for a and b to yield the desired values. This discussion can be extended in an obvious way to include the case of functions of any number of independent variables, so that one can formulate a theorem.

Theorem. *A function of any number of independent variables* $x_1, x_2, \cdots, x_n$ *attains a maximum or a minimum only for those values of the variables* x_i *for which*

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}$$

either vanish simultaneously or cease to exist.

In order to establish a criterion for distinguishing between a maximum and a minimum, note that Taylor's formula gives

$$\begin{aligned} f(a + h, b + k) - f(a, b) &= f_x(a, b)h + f_y(a, b)k \\ &+ \frac{1}{2!}[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2] + R, \end{aligned} \tag{89-3}$$

where

$$R = \frac{1}{3!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^3 f(a + \theta h, b + \theta k).$$

Since the vanishing of the first derivatives is a necessary condition for a maximum or a minimum, (89-3) becomes

$$\begin{aligned} \Delta f &\equiv f(a + h, b + k) - f(a, b) \\ &= \tfrac{1}{2}[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2] + R. \end{aligned} \tag{89-4}$$

From the definition of a maximum it follows that $\Delta f < 0$ for sufficiently small numerical values of h and k, whereas for a minimum it is necessary to demand that $\Delta f > 0$.

In order to simplify the discussion of the determination of the sign of Δf, introduce new independent variables by means of the relations

$$\begin{aligned} h &= r \cos \varphi, \\ k &= r \sin \varphi. \end{aligned}$$

Then (89-4) becomes

$$(89\text{-}5)\quad \Delta f = \frac{r^2}{2}\left(A \cos^2 \varphi + 2B \sin \varphi \cos \varphi + C \sin^2 \varphi + \frac{r}{3}\Phi\right),$$

where

$$\begin{aligned} A &\equiv f_{xx}(a, b), \\ B &\equiv f_{xy}(a, b), \\ C &\equiv f_{yy}(a, b), \end{aligned}$$

and Φ is a function that is bounded for sufficiently small values of r.

If $A \neq 0$, (89-5) can be expressed in the more useful form

$$(89\text{-}6)\quad \Delta f = \frac{r^2}{2}\left[\frac{(A \cos \varphi + B \sin \varphi)^2 + (AC - B^2) \sin^2 \varphi}{A} + \frac{r}{3}\Phi\right].$$

Consider the following possible cases:

(*a*) $AC - B^2 > 0$. In this case the sign of

$$(89\text{-}7)\qquad \frac{(A \cos \varphi + B \sin \varphi)^2 + (AC - B^2) \sin^2 \varphi}{A}$$

is the same as the sign of A. But, since Φ is a bounded function for sufficiently small values of r, the numerical value of $\frac{r}{3}\Phi$ can be made as small as desired by choosing r sufficiently small. Accordingly, the sign of Δf will be determined by that of A. It follows then that the point (a, b) gives a maximum value of the function $f(x, y)$ if $A < 0$, and a minimum value if $A > 0$.

(*b*) $AC - B^2 < 0$. The first term in the numerator of (89-7) vanishes if

$$(89\text{-}8)\qquad \tan \varphi = -\frac{A}{B}.$$

If $\varphi = \varphi_1$ is one of the values satisfying Eq. (89-8), then the numerator of (89-7) will be negative, and by choosing r sufficiently small, the first term in the bracket of (89-6) can be made to dominate the expression. Therefore, for $\varphi = \varphi_1$, Δf will have the sign opposite to that of A. On the other hand, if $\varphi = 0$, the numerator of (89-7) is positive, and hence, for $\varphi = 0$ and sufficiently small values of r the sign of Δf is the same as that of A.

It follows that Δf does not maintain the same sign in the neighborhood of the point (a, b), and consequently, this case gives neither a maximum nor a minimum.

(*c*) $AC - B^2 = 0$. If $AC - B^2$ vanishes, then the numerator of (89-7) is either positive or zero. Therefore, it is necessary to investigate the behavior of the function Φ in order to be able to say anything about the sign of Δf. Since the study of the function Φ is extremely involved and depends upon a very painstaking study of the higher derivatives of the function $f(x, y)$, it will not be pursued here.*

There remains to be considered the case when $A = 0$. But a reference to (89-5) shows that, since

$$2B \sin \varphi \cos \varphi + C \sin^2 \varphi$$

vanishes for some values of φ, this case also leads to the investigation of the function Φ.

It is clear from the foregoing discussion that a maximum or a minimum will surely obtain if $B^2 - AC < 0$, that is, if

$$D \equiv \left(\frac{\partial^2 f}{\partial x\, \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0.$$

A maximum corresponds to the case in which both $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial y^2}$ are negative, and a minimum to the case in which they are both positive.

Example 1. A long piece of tin 12 in. wide is made into a trough by bending up the sides to form equal angles with the base (Fig. 74). Find the amount to be bent up and the angle of inclination of the sides that will make the carrying capacity a maximum.

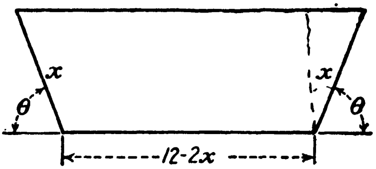

FIG. 74.

The volume will be a maximum if the area of the trapezoidal cross section is a maximum. The area is

$$A = 12x \sin \theta - 2x^2 \sin \theta + x^2 \sin \theta \cos \theta,$$

---

* See GOURSAT, E. Cours d'analyse mathématique, vol. 1, 5th ed., pp. 110–115.

since $12 - 2x$ is the lower base, $12 - 2x + 2x \cos \theta$ is the upper base, and $x \sin \theta$ is the altitude. Then,

$$\begin{aligned}\frac{\partial A}{\partial \theta} &= 12x \cos \theta - 2x^2 \cos \theta + x^2 \cos^2 \theta - x^2 \sin^2 \theta \\ &= x(12 \cos \theta - 2x \cos \theta + x \cos^2 \theta - x \sin^2 \theta)\end{aligned}$$

and

$$\frac{\partial A}{\partial x} = 2 \sin \theta(6 - 2x + x \cos \theta).$$

$\dfrac{\partial A}{\partial x} = 0$ and $\dfrac{\partial A}{\partial \theta} = 0$, if $\sin \theta = 0$ and $x = 0$, which, from physical considerations, cannot give a maximum.

There remain to be satisfied

$$6 - 2x + x \cos \theta = 0$$

and

$$12 \cos \theta - 2x \cos \theta + x \cos^2 \theta - x \sin^2 \theta = 0.$$

Solving the first equation for x and substituting in the second yields, upon simplification,

$$\cos \theta = \tfrac{1}{2} \qquad \text{or} \qquad \theta = 60^\circ, \qquad \text{and} \qquad x = 4.$$

Since physical considerations show that a maximum exists, $x = 4$ and $\theta = 60^\circ$ must give the maximum carrying capacity.

Example 2. Find the maxima and minima of the surface

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2cz.$$

Now,

$$\frac{\partial z}{\partial x} = \frac{1}{c}\frac{x}{a^2}, \qquad \frac{\partial z}{\partial y} = -\frac{1}{c}\frac{y}{b^2},$$

which vanish when $x = y = 0$. But

$$\frac{\partial^2 z}{\partial x^2} = \frac{1}{a^2 c}, \qquad \frac{\partial^2 z}{\partial y^2} = -\frac{1}{b^2 c}, \qquad \frac{\partial^2 z}{\partial x\, \partial y} = 0.$$

Hence, $D = \dfrac{1}{a^2 b^2 c^2}$, and consequently, there is no maximum or minimum at $x = y = 0$. The surface under consideration is a saddle-shaped surface called a *hyperbolic paraboloid*. The points for which the first partial derivatives vanish and $D > 0$ are called *minimax*. The reason for this odd name appears from a consideration of the shape of the hyperbolic paraboloid near the origin of the coordinate system. The reader will benefit from sketching it in the vicinity of (0, 0, 0).

PROBLEMS

1. Divide a into three parts such that their product is a maximum. Test by using the second derivative criterion.

2. Find the volume of the largest rectangular parallelepiped that can be inscribed in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

3. Find the dimensions of the largest rectangular parallelepiped which has three faces in the coordinate planes and one vertex in the plane

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

4. A pentagonal frame is composed of a rectangle surmounted by an isosceles triangle. What are the dimensions for maximum area of the pentagon if the perimeter is given as P?

5. A floating anchorage is designed with a body in the form of a right-circular cylinder with equal ends which are right-circular cones. If the volume is given, find the dimensions giving the minimum surface area.

6. Given n points P_i whose coordinates are (x_i, y_i, z_i), $(i = 1, 2, \cdots, n)$. Show that the coordinates of the point $P(x, y, z)$, such that the sum of the squares of the distances from P to the P_i is a minimum, are given by

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i, \frac{1}{n}\sum_{i=1}^{n} z_i\right).$$

7. Let $z = f(x, y)$ be the equation of a surface which has a minimum at the point (0, 0). Then every plane containing the z-axis will intersect the surface in a curve which has a minimum at (0, 0). The converse of this statement is not true, as can be seen by examining the surface

$$z = (y - x^2)(y - 2x^2)$$

in the neighborhood of the origin. Show that there exists no two-dimensional neighborhood of the point (0, 0) in which the values of this function are all of the same sign.

90. Constrained Maxima and Minima. In a large number of practical and theoretical investigations it is required that a maximum or minimum value of a function be found when the variables are connected by some relation. Thus, it may be required to find a maximum of $u = f(x, y, z)$, where x, y, and z are connected by the relation $\varphi(x, y, z) = 0$. The resulting maximum is called a *constrained maximum.*

The method of obtaining maxima and minima described in the preceding section can be used to solve a problem of constrained maxima and minima as follows: If the constraining relation $\varphi(x, y, z) = 0$ can be solved for one of the variables, say z, in terms of the remaining two variables, and if the resulting expression is substituted for z in $u = f(x, y, z)$, there will be obtained a function $u = F(x, y)$. The values of x and y that yield maxima and minima of u can be found by the methods of Sec. 89. However, the solution of $\varphi(x, y, z) = 0$ for any one of the variables may be extremely difficult, and it is desirable to consider an ingenious device used by Lagrange.

To avoid circumlocution the maximum and minimum values of a function of any number of variables will be called its *extremal values.* It follows from Sec. 89 that the necessary condition for the existence of an extremum of a differentiable function $f(x_1, x_2, \cdots, x_n)$ is the vanishing of the first partial derivatives of the function with respect to the independent variables $x_1, x_2, \cdots, x_n$. Inasmuch as the differential of a function is defined as

$$df \equiv \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n,$$

it is clear that df vanishes for those values of $x_1, x_2, \cdots, x_n$ for which the function has extremal values. Conversely, since the variables x_i are assumed to be independent, the vanishing of the differential is the necessary condition for an extremum.

It is not difficult to see that even when some of the variables are not independent, the vanishing of the total differential is the necessary condition for an extremum. Thus, consider a function

$$u = f(x, y, z), \tag{90-1}$$

where one of the variables, say z, is connected with x and y by some constraining relation

$$\varphi(x, y, z) = 0. \tag{90-2}$$

Regarding x and y as the independent variables, the necessary conditions for an extremum give $\frac{\partial u}{\partial x} = 0$ and $\frac{\partial u}{\partial y} = 0$, or

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0,$$
$$\frac{\partial u}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0.$$

Then the total differential

$$\frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy\right) = 0,$$

and, since the expression in the parenthesis is precisely dz, it follows that

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz = 0. \tag{90-3}$$

The total differential of the constraining relation (90-2) is

$$\frac{\partial \varphi}{\partial x}\,dx + \frac{\partial \varphi}{\partial y}\,dy + \frac{\partial \varphi}{\partial z}\,dz = 0. \tag{90-4}$$

Let this equation be multiplied by some undetermined multiplier λ and then added to (90-3). The result is

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \varphi}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \varphi}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial \varphi}{\partial z}\right)dz = 0.$$

Now if λ is so chosen that

$$\begin{cases} \dfrac{\partial f}{\partial x} + \lambda\dfrac{\partial \varphi}{\partial x} = 0, \\ \dfrac{\partial f}{\partial y} + \lambda\dfrac{\partial \varphi}{\partial y} = 0, \\ \dfrac{\partial f}{\partial z} + \lambda\dfrac{\partial \varphi}{\partial z} = 0, \\ \varphi(x, y, z) = 0, \end{cases} \tag{90-5}$$

then the necessary condition for an extremum of (90-1) will surely be satisfied.

Thus, in order to determine the extremal values of (90-1), it is merely necessary to obtain the solution of the system of equations (90-5) for the four unknowns x, y, z, and λ. The multiplier λ is called a *Lagrangian multiplier.*

Example 1. Find the maximum and the minimum distances from the origin to the curve

$$5x^2 + 6xy + 5y^2 - 8 = 0.$$

The problem here is to determine the extremal values of

$$f(x, y) = x^2 + y^2$$

subject to the condition

$$\varphi(x, y) \equiv 5x^2 + 6xy + 5y^2 - 8 = 0.$$

The equations (90-5) in this case become

$$\begin{aligned} 2x + \lambda(10x + 6y) &= 0, \\ 2y + \lambda(6x + 10y) &= 0, \\ 5x^2 + 6xy + 5y^2 - 8 &= 0. \end{aligned}$$

Multiplying the first of these equations by y and the second by x, and then subtracting gives

$$6\lambda(y^2 - x^2) = 0,$$

so that $y = \pm x$. Substituting these values of y in the third equation gives two equations for the determination of x, namely,

$$2x^2 = 1 \qquad \text{and} \qquad x^2 = 2.$$

The first of these gives $f \equiv x^2 + y^2 = 1$ and the second gives $f \equiv x^2 + y^2 = 4$. Obviously, the first value is a minimum, whereas the second is a maximum. The curve is an ellipse of semiaxes 2 and 1 whose major axis makes an angle of 45° with the x-axis.

Example 2. Find the dimensions of the rectangular box, without a top, of maximum capacity whose surface is 108 sq. in.

The function to be maximized is

$$f(x, y, z) \equiv xyz,$$

subject to the condition

$$xy + 2xz + 2yz = 108. \tag{90-6}$$

The first three of the equations (90-5) become

$$\begin{cases} yz + \lambda(y + 2z) = 0, \\ xz + \lambda(x + 2z) = 0, \\ xy + \lambda(2x + 2y) = 0. \end{cases} \tag{90-7}$$

In order to solve these equations, multiply the first by x, the second by y, and the last by z, and add. There results

$$\lambda(2xy + 4xz + 4yz) + 3xyz = 0,$$

or

$$\lambda(xy + 2xz + 2yz) + \tfrac{3}{2}xyz = 0.$$

Substituting from (90-6) gives

$$108\lambda + \tfrac{3}{2}xyz = 0,$$

or

$$\lambda = -\frac{xyz}{72}.$$

Substituting this value of λ in (90-7) and dividing out common factors gives

$$1 - \frac{x}{72}(y + 2z) = 0,$$

$$1 - \frac{y}{72}(x + 2z) = 0,$$

$$1 - \frac{z}{72}(2x + 2y) = 0.$$

From the first two of these equations it is evident that $x = y$. The substitution of $x = y$ in the third equation gives $z = \dfrac{18}{y}$. Substituting for y and z in the first equation yields $x = 6$. Thus $x = 6$, $y = 6$, and $z = 3$ give the desired dimensions.

PROBLEMS

1. Work Probs. 1, 2, and 3 of Sec. 89 by using Lagrangian multipliers.

2. Prove that the point of intersection of the medians of a triangle possesses the property that the sum of the squares of its distances from the vertices is a minimum.

3. Find the maximum and the minimum of the sum of the angles made by a line from the origin with (*a*) the coordinate axes of a cartesian system; (*b*) the coordinate planes.

4. Find the maximum distance from the origin to (*a*) the folium of Descartes $x^3 + y^3 - 3axy = 0$; (*b*) the ellipse

$$11x^2 + 6xy + 3y^2 - 12x - 12y - 12 = 0.$$

5. Find the shortest distance from the origin to the plane

$$ax + by + cz = d.$$

91. Lagrange's Multipliers. The discussion of the preceding section, showing that the vanishing of the total differential is the necessary condition for an extremum, is perfectly general and can be extended immediately to the function $f(x_1, x_2, \cdots, x_n)$ of n variables, several of which may not be independent.

Thus, consider a function

$$w = f(x, y, z, u), \tag{91-1}$$

and let the constraining relations be

$$\begin{cases} \varphi_1(x, y, z, u) = 0, \\ \varphi_2(x, y, z, u) = 0. \end{cases} \tag{91-2}$$

The relations (91-2) will be thought of as defining the variables z and u as functions of the independent variables x and y. Then the necessary conditions for an extremum are

$$\frac{\partial w}{\partial x} = 0 \quad \text{and} \quad \frac{\partial w}{\partial y} = 0,$$

or

$$\frac{\partial w}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} = 0,$$
$$\frac{\partial w}{\partial y} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} = 0.$$

Multiplying the first of these equations by dx, the second by dy, and then adding,

$$dw = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial u}du = 0. \tag{91-3}$$

The constraining relations (91-2) give

$$\text{(91-4)}\qquad \begin{cases} \dfrac{\partial \varphi_1}{\partial x}\,dx + \dfrac{\partial \varphi_1}{\partial y}\,dy + \dfrac{\partial \varphi_1}{\partial z}\,dz + \dfrac{\partial \varphi_1}{\partial u}\,du = 0, \\ \dfrac{\partial \varphi_2}{\partial x}\,dx + \dfrac{\partial \varphi_2}{\partial y}\,dy + \dfrac{\partial \varphi_2}{\partial z}\,dz + \dfrac{\partial \varphi_2}{\partial u}\,du = 0. \end{cases}$$

Multiplying the first of the equations (91-4) by λ_1 and the second by λ_2, and adding the resulting equations to (91-3),

$$\left(\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial \varphi_1}{\partial x} + \lambda_2 \frac{\partial \varphi_2}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial \varphi_1}{\partial y} + \lambda_2 \frac{\partial \varphi_2}{\partial y}\right) dy$$
$$+ \left(\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial \varphi_1}{\partial z} + \lambda_2 \frac{\partial \varphi_2}{\partial z}\right) dz + \left(\frac{\partial f}{\partial u} + \lambda_1 \frac{\partial \varphi_1}{\partial u} + \lambda_2 \frac{\partial \varphi_2}{\partial u}\right) du = 0.$$

This equation is satisfied if λ_1 and λ_2 are so chosen that

$$\text{(91-5)}\qquad \begin{cases} \dfrac{\partial f}{\partial x} + \lambda_1 \dfrac{\partial \varphi_1}{\partial x} + \lambda_2 \dfrac{\partial \varphi_2}{\partial x} = 0, \\ \dfrac{\partial f}{\partial y} + \lambda_1 \dfrac{\partial \varphi_1}{\partial y} + \lambda_2 \dfrac{\partial \varphi_2}{\partial y} = 0, \\ \dfrac{\partial f}{\partial z} + \lambda_1 \dfrac{\partial \varphi_1}{\partial z} + \lambda_2 \dfrac{\partial \varphi_2}{\partial z} = 0, \\ \dfrac{\partial f}{\partial u} + \lambda_1 \dfrac{\partial \varphi_1}{\partial u} + \lambda_2 \dfrac{\partial \varphi_2}{\partial u} = 0. \end{cases}$$

The four equations (91-5), together with the two constraining relations (91-2), can be used to determine the values of x, y, z, u, λ_1, and λ_2, which give extremal values of (91-1).

The procedure just outlined can be extended in an obvious way to cover the case of more than two constraining relations and more than four variables. This discussion leads to the following rule:

Rule. *In order to determine the extremal values of a function*

$$\text{(91-6)}\qquad f(x_1, x_2, \cdots, x_n)$$

whose variables are subjected to m constraining relations

$$\text{(91-7)}\qquad \varphi_i(x_1, x_2, \cdots, x_n) = 0, \qquad (i = 1, 2, \cdots, m),$$

form the function

$$F = f + \sum_{i=1}^{m} \lambda_i \varphi_i,$$

and determine the parameters λ_i *and the values of* $x_1, x_2, \cdots, x_n$ *from the n equations*

$$\frac{\partial F}{\partial x_j} = 0, \qquad (j = 1, 2, \cdots, n), \tag{91-8}$$

and the m equations (91-7).

As an illustration, consider the problem of determining the maximum and the minimum distances from the origin to the curve of intersection of the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

with the plane

$$Ax + By + Cz = 0.$$

The square of the distance from the origin to any point (x, y, z) is

$$f = x^2 + y^2 + z^2,$$

and it is necessary to find the extremal values of this function when the point (x, y, z) is common to the ellipsoid and the plane. The constraining relations are, therefore,

$$\varphi_1 \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0, \tag{a}$$

and

$$\varphi_2 \equiv Ax + By + Cz = 0. \tag{b}$$

The function $F = f + \lambda_1\varphi_1 + \lambda_2\varphi_2$ is, in this case,

$$F = x^2 + y^2 + z^2 + \lambda_1\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1\right) + 2\lambda_2(Ax + By + Cz),$$

where the factor of 2 is introduced in the last term for convenience. Equations (91-8) then become

$$\begin{cases} x + \lambda_1\dfrac{x}{a^2} + \lambda_2 A = 0, \\ y + \lambda_1\dfrac{y}{b^2} + \lambda_2 B = 0, \\ z + \lambda_1\dfrac{z}{c^2} + \lambda_2 C = 0. \end{cases} \tag{c}$$

These equations, together with (*a*) and (*b*), give five equations for the determination of the five unknowns x, y, z, λ_1, and λ_2. If the first, second, and third of equations (*c*) are multiplied by x, y, and z, respectively, and then added, there results

$$x^2 + y^2 + z^2 + \lambda_1\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) + \lambda_2(Ax + By + Cz) = 0.$$

Making use of (*a*) and (*b*), it is evident that

$$\lambda_1 = -(x^2 + y^2 + z^2) = -f.$$

Setting this value of λ_1 in (*c*) and solving for x, y, and z,

$$x\left(1 - \frac{f}{a^2}\right) + \lambda_2 A = 0, \qquad \text{or} \qquad x = -\frac{\lambda_2 A a^2}{a^2 - f};$$

$$y\left(1 - \frac{f}{b^2}\right) + \lambda_2 B = 0, \qquad \text{or} \qquad y = -\frac{\lambda_2 B b^2}{b^2 - f};$$

$$z\left(1 - \frac{f}{c^2}\right) + \lambda_2 C = 0, \qquad \text{or} \qquad z = -\frac{\lambda_2 C c^2}{c^2 - f}.$$

When these values of x, y, and z are substituted in (*b*) one obtains

$$\frac{A^2a^2}{a^2 - f} + \frac{B^2b^2}{b^2 - f} + \frac{C^2c^2}{c^2 - f} = 0,$$

from which f can be readily determined by solving the quadratic equation in f.

PROBLEMS

1. Find the point P, in the plane of the triangle ABC, for which the sum of the distances from the vertices is a minimum.*

2. Find the triangle of minimum perimeter which can be inscribed in a given triangle.

---

* See E. Goursat's Mathematical Analysis, vol. 1, English ed., p. 130, for a detailed discussion of this interesting problem.

CHAPTER X

IMPROPER INTEGRALS

In defining the definite integral, $\int_a^b f(x)\,dx$, it is assumed that the function $f(x)$ remains bounded in the interval (a, b), and that the end points a and b of the interval are finite. If either (or both) of these assumptions is not satisfied it is necessary to attach a new interpretation to the symbol $\int_a^b f(x)\,dx$. If the function $f(x)$ becomes infinite in the interval $a \le x \le b$, or if the limits of integration become infinite, then the symbol $\int_a^b f(x)\,dx$ is called an improper integral. It will be seen that the improper integrals are defined as limits of certain functions which arise from a consideration of ordinary definite integrals.

92. Integral with Infinite Limit. If the function $f(x)$ is bounded* and integrable for $x \ge a$, then the symbol

$$\int_a^\infty f(x)\,dx \tag{92-1}$$

is defined by the equation

$$\int_a^\infty f(x)\,dx \equiv \lim_{b \to \infty} \int_a^b f(x)\,dx.$$

If this limit exists, then the integral (92-1) is said to converge; otherwise it is called a *divergent integral.* The divergent integral can arise in two ways. It may happen that $\int_a^b f(x)\,dx$ becomes infinite with b, or it may be that this integral oscillates without approaching any limit. Thus,

---

* The definition of an integrable function $f(x)$ (Sec. 35) demands that $f(x)$ be bounded; hence, the statement that the function is *bounded and integrable* is redundant. However, some authors extend the class of functions integrable in the sense of Riemann to include those unbounded functions whose improper integrals exist (see Sec. 94). The word *bounded* is inserted here to preclude the possibility of interpreting the term *integrable function* in the extended sense.

$$\int_a^\infty \frac{dx}{\sqrt{x}} = \lim_{b\to\infty} \int_a^b \frac{dx}{\sqrt{x}} = \lim_{b\to\infty} (2\sqrt{b} - 2\sqrt{a}).$$

The expression within the parentheses becomes infinite with b. On the other hand, for

$$\int_a^\infty \sin x \, dx = \lim_{b\to\infty} \int_a^b \sin x \, dx = \lim_{b\to\infty} (\cos a - \cos b)$$

the expression in parentheses oscillates as b is increased and does not approach any limit. Both of these integrals are called divergent. However,

$$\int_a^\infty \frac{x\,dx}{(x^2+1)^3} = \lim_{b\to\infty} \int_a^b \frac{x\,dx}{(x^2+1)^3}$$
$$= \lim_{b\to\infty} \frac{1}{4}\left[\frac{1}{(a^2+1)^2} - \frac{1}{(b^2+1)^2}\right]$$

converges to the value

$$\frac{1}{4(a^2+1)^2}.$$

Similarly,

$$\int_{-\infty}^b f(x)\,dx$$

is defined as the limit of the function $\int_a^b f(x)\,dx$ when a becomes negatively infinite, so that

$$\int_{-\infty}^b f(x)\,dx \equiv \lim_{a\to-\infty} \int_a^b f(x)\,dx.$$

It may happen that the integral $\int_a^b f(x)\,dx$ of a bounded function $f(x)$ exists for all finite values of a and b; in that case the symbol

$$\int_{-\infty}^\infty f(x)\,dx$$

is understood to mean

$$\int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx, \tag{92-2}$$

where c is any real number.

It is clear that a study of an improper integral with a continuous integrand $f(x)$ reduces to a study of the behavior of the limit of the function $F(x)$ as $x \to \infty$, where $F(x) \equiv \int_a^x f(x)\,dx$. The conditions under which a function $F(x)$ converges to a limit have been discussed in Chap. I, and a direct application of the fundamental principle of convergence to a limit furnishes the following theorem.

Theorem 1. *A necessary and sufficient condition that* $\int_a^\infty f(x)\,dx$ *be convergent is that, for any arbitrarily small positive number* ϵ, *there exist a number* N *such that*

$$\left|\int_p^q f(x)\,dx\right| < \epsilon$$

for all values of p *and* q *which both exceed* N.

However, the application of Theorem 1 to the determination of the convergence of a particular integral is likely to be difficult. For this reason some simple practical tests are developed with the aid of Theorem 1 and the following theorem.

Theorem 2. *If* $f(x)$ *is continuous for* $x \geq a$ *and if* $\int_a^\infty |f(x)|\,dx$ *exists, then* $\int_a^\infty f(x)\,dx$ *also exists. The function* $f(x)$ *is then said to be absolutely integrable over the infinite interval.*

In order to prove this statement, write

$$f(x) \equiv [f(x) + |f(x)|] - |f(x)|.$$

Then

$$(92\text{-}3)\qquad \int_a^b f(x)\,dx = \int_a^b [f(x) + |f(x)|]\,dx - \int_a^b |f(x)|\,dx.$$

By hypothesis the second integral in the right-hand member of (92-3) converges when $b \to \infty$. But

$$0 \leq f(x) + |f(x)| \leq 2|f(x)|;$$

hence

$$0 \leq \int_a^b [f(x) + |f(x)|]\,dx \leq \int_a^b 2|f(x)|\,dx.$$

Since the integral $\int_a^b 2|f(x)|\,dx$ is convergent as $b \to \infty$, it follows that the first integral in the right-hand member of (92-3) also converges.

The converse of Theorem 2 is not true. The integral of $f(x)$ may converge over an infinite interval when the integral of

$|f(x)|$ does not converge. This corresponds to the situation in infinite series when a series converges conditionally.

As an instance of such behavior consider the integral*

$$\int_0^\infty \frac{\sin x}{x}\, dx, \tag{92-4}$$

which is convergent but not absolutely convergent. The convergence of the integral (92-4) will be established with the aid of Theorem 1.

It follows from the second mean-value theorem for integrals (Sec. 37) that

$$\int_p^q \frac{\sin x}{x}\, dx = \frac{1}{p}\int_p^\xi \sin x\, dx + \frac{1}{q}\int_\xi^q \sin x\, dx,$$

where $0 < p \leq \xi \leq q$. But

$$\left|\int_p^\xi \sin x\, dx\right| \leq 2 \qquad \text{and} \qquad \left|\int_\xi^q \sin x\, dx\right| \leq 2.$$

Consequently,

$$\begin{aligned}\left|\int_p^q \frac{\sin x}{x}\, dx\right| &\leq 2\left(\frac{1}{p} + \frac{1}{q}\right)\\ &< 2\left(\frac{1}{p} + \frac{1}{p}\right)\\ &= \frac{4}{p}.\end{aligned}$$

Now if any $\epsilon > 0$ is prescribed, the inequality

$$\left|\int_p^q \frac{\sin x}{x}\, dx\right| < \epsilon$$

will surely be satisfied for all values of p and q that are greater than $N = \frac{4}{\epsilon}\cdot$ Thus, the integral (92-4) is convergent.†

---

* It is customary to define the value of $\frac{\sin x}{x}$ at $x = 0$ as unity, so that $\frac{\sin x}{x}$ is a continuous function.

† It will be shown in Sec. 97 that $\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}\cdot$

In order to prove that

$$\int_0^\infty \left|\frac{\sin x}{x}\right| dx \tag{92-5}$$

diverges, note that for any positive integral value of k

$$\int_0^{(k+1)\pi} \left|\frac{\sin x}{x}\right| dx = \int_0^{\pi} \left|\frac{\sin x}{x}\right| dx + \int_\pi^{2\pi} \left|\frac{\sin x}{x}\right| dx + \cdots + \int_{k\pi}^{(k+1)\pi} \left|\frac{\sin x}{x}\right| dx$$

$$= \sum_{n=0}^{k} \int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| dx.$$

Now, for $n \geq 0$,

$$\int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| dx \geq \int_{n\pi}^{(n+1)\pi} \frac{|\sin x|}{(n+1)\pi} dx = \frac{2}{(n+1)\pi}, \tag{92-6}$$

since the minimum value of $\dfrac{1}{x}$ in the interval $[n\pi, (n+1)\pi]$ is $\dfrac{1}{(n+1)\pi}$. It follows that

$$\sum_{n=0}^{k} \int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| dx \geq \frac{2}{\pi}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k+1}\right)$$

and, since the series in the right-hand member of this inequality diverges as $k \to \infty$, the integral (92-5) is also divergent.

It may be remarked that the function $f(x)$ need not approach zero as $x \to \infty$ in order that the integral $\int_0^\infty f(x)\, dx$ be convergent. Consider, for example, the integral

$$\int_0^\infty \sin x^2\, dx,$$

and assume that it can be written as the sum of a series of integrals, namely,

$$\int_0^\infty \sin x^2\, dx = \sum_{n=0}^{\infty} \int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} \sin x^2\, dx. \tag{92-7}$$

The series in the right-hand member of (92-7) is readily shown to be convergent, but the integrand of (92-7) obviously does not approach zero as $x \to \infty$.

In order to prove the convergence of the series, let the variable x be changed by the substitution

$$x = \sqrt{z + n\pi}.$$

Then

$$\int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} \sin x^2\,dx = \frac{1}{2}\int_0^\pi \frac{\sin(z + n\pi)}{\sqrt{z + n\pi}}\,dz = \frac{1}{2}\int_0^\pi \frac{(-1)^n \sin z}{\sqrt{z + n\pi}}\,dz.$$

It is clear from this form of the general term of the series that the series in the right-hand member of (92-7) is alternating and decreasing in such a way that the limit of the nth term as $n \to \infty$ is zero. Consequently, the series and, hence, the integral (92-7) converge.

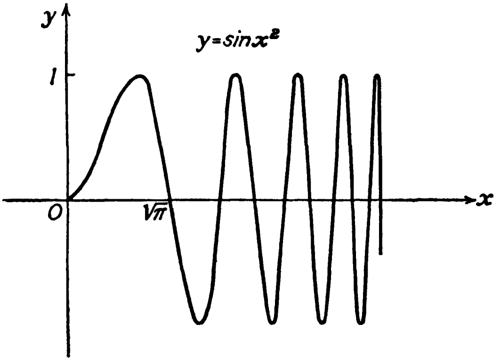

FIG. 75.

The question of the legitimacy of the infinite series representation (92-7) of the integral $\int_0^\infty \sin x^2\,dx$ will be settled in Sec. 96, but it is not at all necessary to resort to this method of establishing the convergence. In fact, a substitution of $x^2 = y$ leads to a consideration of the integral $\int_0^\infty \frac{\sin y}{\sqrt{y}}\,dy$, which is readily shown to be convergent.*

---

* See the illustration at the end of Sec. 93 and Example 2 of Sec. 95.

The reason for the convergence of this integral may be seen from geometrical considerations (Fig. 75). The areas bounded by the successive arches of $y = \sin x^2$ decrease with the increase in n, because the difference of consecutive roots, $\sqrt{(n+1)\pi} - \sqrt{n\pi}$, of $\sin x^2$ tends to zero when $n \to \infty$, while the maximum height of each arch always remains equal to unity.

93. Tests for Convergence of Integrals with Infinite Limits. This section contains some practical tests for determining the convergence of the integral $\int_a^\infty f(x)\,dx$, where $f(x)$ is a continuous function for all values of $x \geq a$. It will be assumed throughout this section that a is a positive number. This entails no loss of generality, since the integral $\int_{a_0}^\infty f(x)\,dx$, where the lower limit a_0 is negative, can be written as the sum of two integrals, namely

$$\int_{a_0}^\infty f(x)\,dx = \int_{a_0}^a f(x)\,dx + \int_a^\infty f(x)\,dx,$$

where a is a positive number.

Theorem 1. *If $f(x)$ is continuous for $x \geq a$ and there exists a positive number A such that*

$$|f(x)x^k| < A \qquad \textit{when} \qquad x \geq a,$$

then $\int_a^\infty f(x)\,dx$ converges absolutely if $k > 1$. If $|f(x)x^k| \geq A$ and $k \leq 1$, then the integral $\int_a^\infty f(x)\,dx$ diverges.

The proof of the theorem is simple. For, if

$$|f(x)| < \frac{A}{x^k},$$

then

$$\int_a^\infty |f(x)|\,dx < A\int_a^\infty \frac{1}{x^k}\,dx. \tag{93-1}$$

But

$$\int_a^b \frac{1}{x^k}\,dx = \frac{1}{1-k}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right), \qquad \text{if} \qquad k \neq 1, \tag{93-2}$$

and, if $k > 1$, the right-hand member of (93-2) approaches the limit $\dfrac{1}{(k-1)a^{k-1}}$ when $b \to \infty$. Thus, the integral on the right of the inequality (93-1) exists, and, since $\int_a^x |f(x)|\,dx$ is an increasing function of x, the integral $\int_a^\infty |f(x)|\,dx$ exists. It follows from Theorem 2, Sec. 92, that $\int_a^\infty f(x)\,dx$ exists.

The proof of the second part of the theorem is just as easy. For

$$|f(x)| \geq \frac{A}{x^k} \qquad \text{when} \qquad x \geqq a.$$

Therefore,

$$\int_a^b |f(x)|\,dx \geq A \int_a^b \frac{1}{x^k}\,dx.$$

But if $|f(x)| \geq \frac{A}{x^k}$, then $f(x)$ cannot change sign for $x \geq a$, since $f(x)$ is a continuous function. Suppose that $f(x) > 0$; then

$$\int_a^b |f(x)|\,dx = \int_a^b f(x)\,dx \geq A \int_a^b \frac{1}{x^k}\,dx.$$

But

$$\begin{aligned} \int_a^b \frac{1}{x^k}\,dx &= \frac{1}{1-k}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right), &\quad \text{if} \quad k < 1; \\ &= \log b - \log a, &\quad \text{if} \quad k = 1, \end{aligned}$$

which diverges as $b \to \infty$. Consequently, $\int_a^\infty f(x)\,dx$ diverges. If $f(x) < 0$ for $x \geq a$, then

$$\int_a^b |f(x)|\,dx = -\int_a^b f(x)\,dx,$$

and the assertion that $\int_a^b f(x)\,dx$ diverges follows immediately from the proof for the case when $f(x) > 0$.

This theorem is capable of a simple geometrical interpretation. If $|f(x)| < \frac{A}{x^k}$, then the graph of the function $y = f(x)$ lies between the graphs of $y = \pm\frac{A}{x^k}$ (Fig. 76). Consequently, the area under the curve $y = \frac{A}{x^k}$, bounded by the ordinates

$x = a$ and $x = b$ and the x-axis, is greater than the area under the curve $y = |f(x)|$, so that

$$\int_a^b |f(x)|\,dx < \int_a^b \frac{A}{x^k}\,dx.$$

If the integral on the right of this inequality is convergent, then the integral on the left will surely converge.

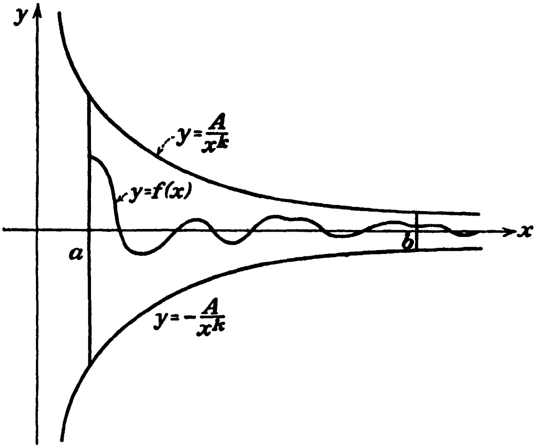

FIG. 76.

If, on the other hand,

$$f(x) \geq \frac{A}{x^k} \qquad \text{for} \qquad x \geq a,$$

then the area under the curve $y = f(x)$ will not be less than that under $y = \frac{A}{x^k}$, so that

$$\int_a^b f(x)\,dx \geq \int_a^b \frac{A}{x^k}\,dx.$$

A particularly useful form of Theorem 1 appears when $f(x)$ is such a function that $\lim\limits_{x\to\infty} x^k f(x)$ exists. In such a case Theorem 1 can be phrased as follows:

Corollary. *If* $\lim\limits_{x\to\infty} x^k f(x) = L$, *where* $k > 1$, *then* $\int_a^\infty f(x)\,dx$ *converges absolutely.*

If $\lim\limits_{x\to\infty} x^k f(x) = L \neq 0$ *and* $k \leq 1$, *then* $\int_a^\infty f(x)\,dx$ *diverges.*

As an illustration of the application of this corollary consider

$$\int_0^\infty \frac{dx}{(1+x^6)^{1/3}}.$$

Since

$$\lim_{x\to\infty} x^2 \frac{1}{(1+x^6)^{1/3}} = \lim_{x\to\infty} \frac{1}{\left(\frac{1}{x^6}+1\right)^{1/3}} = 1,$$

it follows that the given integral is convergent.

The reader will have no difficulty in deducing, with the aid of the corollary, the following useful theorem.

Theorem 2. *If* $f(x) = \frac{P_1(x)}{P_2(x)}$, *where* $P_1(x)$ *and* $P_2(x)$ *are polynomials in* x *and* $P_2(x) \neq 0$ *for* $x \geq a$, *then the necessary and sufficient condition that* $\int_a^\infty f(x)\,dx$ *converge is that the degree of* $P_2(x)$ *exceed the degree of* $P_1(x)$ *by at least two.*

Theorem 3. *If* $f(x)$ *is continuous for* $x \geq a$ *and there exists a positive number* A *such that* $|e^{kx}f(x)| < A$ *when* $x \geq a$, *then* $\int_a^\infty f(x)\,dx$ *converges absolutely if* $k > 0$. *If* $|e^{kx}f(x)| \geq A$, *then* $\int_a^\infty f(x)\,dx$ *diverges if* $k \leq 0$.

The proof of this theorem follows by using the argument employed in establishing Theorem 1, and by noting that

$$\begin{aligned} \int_a^b e^{-kx}\,dx &= -\frac{1}{k}(e^{-kb} - e^{-ka}), &&\text{if} \quad k \neq 0,\\ &= b - a, &&\text{if} \quad k = 0. \end{aligned}$$

Corollary. *If* $\lim\limits_{x\to\infty} f(x)e^{kx} = L$ *and* $k > 0$, *then* $\int_a^\infty f(x)\,dx$ *converges absolutely.*

If $\lim\limits_{x\to\infty} f(x)e^{kx} = L \neq 0$ *and* $k \leq 0$, *then* $\int_a^\infty f(x)\,dx$ *diverges.*

As an illustration consider

$$\int_0^\infty e^{-x^2}\,dx.$$

Since

$$\lim_{x\to\infty} e^x e^{-x^2} = \lim_{x\to\infty} e^{x(1-x)} = 0,$$

it follows that the given integral is convergent.

Theorem 4. *If* $\varphi(x)$ *is a monotone function which approaches a finite limit as* x *increases indefinitely, then*

$$\int_a^\infty f(x)\,\varphi(x)\,dx$$

will converge if $\int_a^\infty f(x)\,dx$ *converges.*

The proof of this theorem follows from Theorem 1, Sec. 92, with the aid of the second mean-value theorem for integrals (Sec. 37). The convergence of $\int_a^\infty f(x)\,dx$ requires that there exist a number N such that when p and q are any numbers greater than N, then

$$\left|\int_p^q f(x)\,dx\right| < \epsilon \equiv \frac{\epsilon'}{2k}.$$

But

$$\left|\int_p^q \varphi(x)f(x)\,dx\right| = \left|\varphi(p)\int_p^\xi f(x)\,dx + \varphi(q)\int_\xi^q f(x)\,dx\right|$$
$$\leq \left|\varphi(p)\right|\cdot\left|\int_p^\xi f(x)\,dx\right| + \left|\varphi(q)\right|\cdot\left|\int_\xi^q f(x)\,dx\right|,$$

where p, ξ, and q are all greater than N. Then, if L is the limit approached by $\varphi(x)$ as x increases, and if k is the larger of $|L|$ and $|\varphi(a)|$, it is evident that

$$\left|\varphi(p)\right|\cdot\left|\int_p^\xi f(x)\,dx\right| + \left|\varphi(q)\right|\cdot\left|\int_\xi^q f(x)\,dx\right| < 2\epsilon k = \epsilon'.$$

Therefore,

$$\left|\int_p^q \varphi(x)f(x)\,dx\right| < \epsilon',$$

if p and q are any numbers greater than N.

By a similar argument it can be shown that if $\varphi(x)$ is a monotone function which approaches zero as x increases indefinitely, then $\int_a^\infty \varphi(x)f(x)\,dx$ converges if $\int_a^b f(x)\,dx$ remains bounded as b increases indefinitely. It should be noticed that this statement does not require that $\int_a^b f(x)\,dx$ converge to a limit as b is increased.

For example, the integral

$$\int_1^\infty \frac{\sin x}{x^k}\,dx,$$

where $k > 0$, can be written as

$$\int_1^\infty \frac{1}{x^k}\sin x\,dx.$$

Since $\frac{1}{x^k}$ is a monotone function which approaches zero as x increases, and since $\int_1^b \sin x\,dx = \cos 1 - \cos b$ which oscillates as b increases, it follows that the given integral is convergent. The convergence of this integral was discussed in Sec. 92 for the case in which $k = 1$.

As another example, consider

$$\int_0^\infty e^{-\left(x-\frac{a}{x}\right)^2}\,dx.$$

This integral can be written as

$$\int_0^\infty e^{-x^2+2a-\frac{a^2}{x^2}}\,dx = e^{2a}\int_0^\infty e^{-x^2}e^{-\frac{a^2}{x^2}}\,dx.$$

Now $e^{-\frac{a^2}{x^2}}$ is a monotone-increasing function which approaches the limit 1 as x increases, and $\int_0^\infty e^{-x^2}\,dx$ was shown to converge. Hence,

$$\int_0^\infty e^{-\left(x-\frac{a}{x}\right)^2}\,dx$$

is convergent.

PROBLEMS

Test the following integrals for convergence or divergence:

1. $\displaystyle\int_1^\infty \frac{\cos x}{x^k}\,dx.$
2. $\displaystyle\int_1^\infty \frac{dx}{x\sqrt{1+x^2}}.$
3. $\displaystyle\int_0^\infty \frac{x\,dx}{\sqrt{2+x^3}}.$
4. $\displaystyle\int_0^\infty e^{-x^2}\sin x\,dx.$
5. $\displaystyle\int_2^\infty \frac{e^{\frac{1}{x}}}{(1+x)^2}\,dx.$
6. $\displaystyle\int_1^\infty \frac{e^{-x}\sin x}{x}\,dx.$
7. $\displaystyle\int_1^\infty \frac{\sin x}{x^k}\,dx.$
8. $\displaystyle\int_0^\infty \frac{x\sin x}{1+x^2}\,dx.$
9. $\displaystyle\int_1^\infty \frac{\tan^{-1} x}{1+x^2}\,dx.$
10. $\displaystyle\int_0^\infty \frac{\sin x}{1+x}\,dx.$
11. $\displaystyle\int_0^\infty \frac{dx}{\sqrt{1+x^4}}.$
12. $\displaystyle\int_0^\infty \frac{dx}{1+x^3}.$

94. Integrals in Which the Integrand Becomes Infinite. If $f(x)$ becomes infinite for some value of $x = c$, where $a < c < b$, then, by definition,

$$\int_a^b f(x)\,dx \equiv \lim_{\epsilon_1 \to 0} \int_a^{c-\epsilon_1} f(x)\,dx + \lim_{\epsilon_2 \to 0} \int_{c+\epsilon_2}^b f(x)\,dx, \tag{94-1}$$

and the integral is said to exist if the limits of the two integrals on the right-hand side of (94-1) exist. In case $f(x)$ becomes infinite at $x = a$, then, by definition,

$$\int_a^b f(x)\,dx = \lim_{\epsilon \to 0} \int_{a+\epsilon}^b f(x)\,dx,$$

and if $f(x)$ becomes infinite at $x = b$,

$$\int_a^b f(x)\,dx = \lim_{\epsilon \to 0} \int_a^{b-\epsilon} f(x)\,dx.$$

It may be observed that, since

$$\lim_{\epsilon \to 0} \int_{a+\epsilon}^b f(x)\,dx = \lim_{\epsilon \to 0} [F(b) - F(a + \epsilon)],$$

where $F(x)$ is the indefinite integral of $f(x)$, the integral converges if $\lim\limits_{\epsilon \to 0} F(a + \epsilon)$ exists. If the indefinite integral can be found, the evaluation of the definite integral is accomplished in the process of determining its convergence.

It is important to note that the variables ϵ_1 and ϵ_2 in the definition (94-1) are assumed to approach zero independently of one another. The same variable ϵ is used occasionally in the definition, so that

$$\int_a^b f(x)\,dx \equiv \lim_{\epsilon \to 0} \left(\int_a^{c-\epsilon} f(x)\,dx + \int_{c+\epsilon}^b f(x)\,dx \right). \tag{94-2}$$

But the limit in (94-2) may exist when the limits in (94-1) do not exist. If the limit in (94-2) exists and is equal to L, then the value L is called *Cauchy's principal value of the integral* $\int_a^b f(x)\,dx$.

If $f(x)$ becomes infinite at both end points of the interval but remains bounded within (a, b), then the integral of $f(x)$ is defined by the equation

$$\int_a^b f(x)\,dx \equiv \lim_{\epsilon_1 \to 0} \int_{a+\epsilon_1}^c f(x)\,dx + \lim_{\epsilon_2 \to 0} \int_c^{b-\epsilon_2} f(x)\,dx,$$

where each of the limits on the right must exist if the integral $\int_a^b f(x)\,dx$ is to converge.

The definitions given above can be extended in an obvious way to include the case where $f(x)$ has any finite number of discontinuities in the interval (a, b).

Example 1.

$$\int_0^1 \frac{dx}{x^2} = \lim_{\epsilon \to 0} \int_\epsilon^1 \frac{dx}{x^2} = \lim_{\epsilon \to 0} \left(\frac{1}{\epsilon} - 1\right).$$

Hence, the integral does not exist.

Example 2.

$$\int_0^2 \frac{dx}{\sqrt{2 - x}} = \lim_{\epsilon \to 0} \int_0^{2-\epsilon} \frac{dx}{\sqrt{2 - x}} = \lim_{\epsilon \to 0} \left(-2\sqrt{\epsilon} + 2\sqrt{2}\right),$$

and thus the integral converges to $2\sqrt{2}$.

Example 3.

$$\begin{aligned} \int_0^3 \frac{x\,dx}{(x^2 - 1)^{2/3}} &= \lim_{\epsilon_1 \to 0} \int_0^{1-\epsilon_1} \frac{x\,dx}{(x^2 - 1)^{2/3}} + \lim_{\epsilon_2 \to 0} \int_{1+\epsilon_2}^3 \frac{x\,dx}{(x^2 - 1)^{2/3}} \\ &= \lim_{\epsilon_1 \to 0} \left[\frac{3}{2}(\epsilon_1^2 - 2\epsilon_1)^{1/3} + \frac{3}{2}\right] \\ &\qquad + \lim_{\epsilon_2 \to 0} \left[\frac{3}{2}(2) - \frac{3}{2}(\epsilon_2^2 + 2\epsilon_2)^{1/3}\right]. \end{aligned}$$

Therefore the given integral is convergent to $9/2$.

Example 4.

$$\begin{aligned} \int_{-1}^1 \frac{dx}{x^3} &= \lim_{\epsilon_1 \to 0} \int_{-1}^{-\epsilon_1} \frac{dx}{x^3} + \lim_{\epsilon_2 \to 0} \int_{\epsilon_2}^1 \frac{dx}{x^3} \\ &= -\frac{1}{2}\left[\lim_{\epsilon_1 \to 0} \left(\frac{1}{\epsilon_1^2} - 1\right) + \lim_{\epsilon_2 \to 0} \left(1 - \frac{1}{\epsilon_2^2}\right)\right], \end{aligned}$$

and therefore the given integral does not converge. But if $\epsilon_2 = \epsilon_1 = \epsilon$, then

$$\begin{aligned} \int_{-1}^1 \frac{dx}{x^3} &= \lim_{\epsilon \to 0} \left(\int_{-1}^{\epsilon} \frac{dx}{x^3} + \int_\epsilon^1 \frac{dx}{x^3}\right) \\ &= \lim_{\epsilon \to 0} \left[-\frac{1}{2}\left(\frac{1}{\epsilon^2} - 1 + 1 - \frac{1}{\epsilon^2}\right)\right] = 0. \end{aligned}$$

Hence, the principal value of $\int_{-1}^1 \frac{dx}{x^3}$ is equal to zero.

It is clear from the foregoing that the determination of the convergence of improper integrals, with discontinuous integrands, is reduced to the study of the behavior of the limit of the function $F(\epsilon)$ as $\epsilon \to 0$, where $F(\epsilon)$ is of the form

$$F(\epsilon) = \int_{a+\epsilon}^{b} f(x)\,dx.$$

From the definition of the limit of a function (Sec. 8) and from the fundamental principle of convergence (Sec. 5), it follows that:

The necessary and sufficient condition that $\lim_{\epsilon \to 0+} F(\epsilon)$ *exist is that, for any preassigned positive number* η, *one can find a number* δ *such that*

$$|F(\epsilon_2) - F(\epsilon_1)| < \eta$$

whenever $0 < \epsilon_2 < \epsilon_1 < \delta$.

Recalling the definition of $F(\epsilon)$ one can state the following theorem:

Theorem. *A necessary and sufficient condition for the convergence of the integral*

$$\int_a^b f(x)\,dx$$

whose integrand becomes infinite at $x = a$ *is that for any preassigned number* $\eta > 0$, *one can find a positive number* δ *such that*

$$\left|\int_{a+\epsilon_2}^{a+\epsilon_1} f(x)\,dx\right| < \eta$$

whenever $0 < \epsilon_2 < \epsilon_1 < \delta$.

It follows from this theorem that if $\int_a^b |f(x)|\,dx$ is convergent, then $\int_a^b f(x)\,dx$ will surely converge. The converse of this statement, however, is not true. For, consider the integral

$$\int_0^1 \frac{\sin \frac{1}{x}}{x}\,dx.$$

Setting $x = \frac{1}{y}$ gives the integral

$$\int_1^\infty \frac{\sin y}{y}\,dy,$$

which was shown to be convergent, but not absolutely convergent (see Sec. 92).

95. Tests for Convergence of Integrals Whose Integrands Become Infinite. Corresponding to the tests given in Sec. 93 for the convergence of integrals with infinite limits, one can develop tests for the convergence of integrals whose integrands become infinite in the interval of integration. One of these tests is given next.

Theorem. *If* $f(x)$ *is continuous for* $a < x \leq b$ *and there exists a positive number* A *such that* $|f(x)(x - a)^k| < A$ *for* $0 < k < 1$ *and* $a < x \leq b$, *then* $\int_a^b f(x)\,dx$ *is absolutely convergent.*

If $|f(x)(x - a)^k| > A$ *for* $k \geq 1$ *and* $a < x \leq b$, *then* $\int_a^b f(x)\,dx$ *is divergent.*

The proof of this theorem is entirely similar to that of Theorem 1, Sec. 93. It is only necessary to observe that

$$\int_a^b \frac{dx}{(x - a)^k}$$

exists if $k < 1$ and does not exist if $k \geq 1$.

Corollary. *If* $\lim_{x \to a+} (x - a)^k f(x) = L$, *then* $\int_a^b f(x)\,dx$ *converges absolutely if* $0 < k < 1$.

If $\lim_{x \to a+} (x - a)^k f(x) = L \neq 0$, *then* $\int_a^b f(x)\,dx$ *diverges if* $k \geq 1$.

Example 1. Consider

$$\int_1^2 \frac{dx}{\sqrt{x^3 - 1}}.$$

Since

$$\lim_{x \to 1+} (x - 1)^{1/2} \frac{1}{(x^3 - 1)^{1/2}} = \lim_{x \to 1+} \frac{1}{\sqrt{x^2 + x + 1}} = \frac{1}{\sqrt{3}},$$

it follows that the given integral is convergent.

Example 2. Consider

$$\int_0^1 \frac{\sin x}{x^n}\,dx, \qquad n > 0.$$

Now

$$\frac{\sin x}{x^n} = \frac{\sin x}{x} \cdot \frac{1}{x^{n-1}}.$$

But, since $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$, it follows that

$$\begin{aligned}\lim_{x\to 0} x^k \frac{\sin x}{x^n} &= \lim_{x\to 0} \left(x^k \cdot \frac{\sin x}{x} \cdot \frac{1}{x^{n-1}}\right)\\ &= \lim_{x\to 0} x^{k-n+1}.\end{aligned}$$

This limit will exist provided that $n - 1 \leq k$. Thus, if $n < 2$, the integral is convergent, and if $n \geq 2$, the integral is divergent.

Example 3. Consider

$$\int_0^{\frac{\pi}{2}} \log \sin x \, dx.$$

Note that

$$\begin{aligned}\log \sin x &= \log \left(x \cdot \frac{\sin x}{x}\right)\\ &= \log x + \log \frac{\sin x}{x}.\end{aligned}$$

But

$$\lim_{x\to 0} x^k \log \sin x = \lim_{x\to 0} \left(x^k \log x + x^k \log \frac{\sin x}{x}\right).$$

Since $\lim\limits_{x\to 0} x^k \log x = 0$ if $k > 0$, and $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$, it is evident that $\lim\limits_{x\to 0} x^k \log \sin x = 0$, when $k > 0$. Consequently the integral converges.

PROBLEMS

1. Test the following integrals for convergence or divergence:

(*a*) $\displaystyle\int_0^{\pi} \frac{\cos x}{\sqrt{x}}\, dx$;

(*b*) $\displaystyle\int_0^1 \frac{x^{\alpha-1}}{\sqrt{1-x}}\, dx$;

(*c*) $\displaystyle\int_0^2 \frac{dx}{\sqrt{2x - x^2}}$;

(*d*) $\displaystyle\int_1^2 \frac{dx}{x^3 - 1}$;

(*e*) $\displaystyle\int_0^1 \frac{dx}{\sqrt{x(x-1)}}$;

(*f*) $\displaystyle\int_0^1 \frac{dx}{x(x+1)}$;

(*g*) $\displaystyle\int_0^1 \frac{dx}{(x+1)\sqrt{x}}$;

(*h*) $\displaystyle\int_0^1 \frac{x^{n-1}}{x-1}\, dx$;

(*i*) $\displaystyle\int_0^1 x^n \log x \, dx$;

(*j*) $\displaystyle\int_0^a \frac{\log x}{1-x}\, dx, \qquad a < 1.$

2. Prove the convergence of the integrals

$$(a)\ E = \int_0^a \frac{\sqrt{a^2 - e^2x^2}}{\sqrt{a^2 - x^2}}\, dx, \qquad 0 < e < 1;$$

$$(b)\ F = \int_0^b \frac{dx}{\sqrt{(1 - x^2)(1 - e^2x^2)}}, \qquad b > \frac{1}{e}, \qquad 0 < e < 1.$$

These integrals are known as the elliptic integrals of the second and first kinds, respectively.

3. Prove the convergence of the integrals

$$(a)\ \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - e^2 \sin^2 \theta}}, \qquad 0 < e < 1;$$

$$(b)\ \int_0^{\frac{\pi}{2}} \sqrt{1 - e^2 \sin^2 \theta}\, d\theta, \qquad 0 < e < 1.$$

Set $x = \sin \theta$, and use the results of Prob. 2.

4. If $P(x)$ and $Q(x)$ are relatively prime polynomials, show that $\int_a^b \frac{P(x)}{Q(x)}\, dx$ diverges whenever a root of $Q(x)$ falls in the interval $a \leq x \leq b$.

96. Operations with Improper Integrals. Inasmuch as the improper integrals are defined as the limits of certain functions, it is natural to inquire whether it is permissible to perform the same operations on improper integrals as are familiar in the case of finite integrals. It is clear from a consideration of the examples in Sec. 92 that the behavior of some improper integrals is closely analogous to that of infinite series, and this fact alone is sufficient to indicate the need of caution.

It is true that the equation

$$\int_a^b cf(x)\, dx = c\int_a^b f(x)\, dx,$$

when c is a constant, is valid, even when the integrals are improper, since in applying the limit process, the constant c can be placed either inside or outside the limit sign.

The equation

$$(96\text{-}1)\qquad \int_a^b [f_1(x) + f_2(x)]\, dx = \int_a^b f_1(x)\, dx + \int_a^b f_2(x)\, dx$$

however, is not always true. It may be that $\varphi(x) = f_1(x) + f_2(x)$

is a function whose integral converges, whereas the integrals of $f_1(x)$ and $f_2(x)$ are divergent. Thus,

$$\int_2^\infty \left(-\frac{1}{x}\right) dx$$

is divergent, as is

$$\int_2^\infty \frac{dx}{x-1},$$

but

$$\int_2^\infty \left(\frac{1}{x-1} - \frac{1}{x}\right) dx$$

is convergent. However, if any two of the integrals in (96-1) converge, then the third integral also converges, and Eq. (96-1) is true. Similarly, if any two of the expressions in the formula for integration by parts,

$$\int_a^b f_1(x)f_2'(x)\,dx = f_1(x)f_2(x)\Big|_a^b - \int_a^b f_1'(x)f_2(x)\,dx, \tag{96-2}$$

exist, then the third expression also exists and the relation (96-2) is valid.

Matters are considerably more involved when one considers the differentiation and integration of improper integrals with respect to a parameter, and it is necessary to impose more severe restrictions on the integrands than was the case with the ordinary integrals. Thus, consider an integral containing a parameter, namely,

$$F(\alpha) = \int_0^\infty \frac{\sin \alpha x}{x}\,dx, \qquad \text{where} \qquad \alpha > 0. \tag{96-3}$$

If the differentiation under the integral sign be performed formally there results the integral

$$\int_0^\infty \cos \alpha x\,dx,$$

which is divergent and consequently does not represent the function $F'(\alpha)$.*

---

* It may be remarked that the function defined by the integral (96-3) is a constant. For. setting $z = \alpha x$ gives $\int_0^\infty \frac{\sin z}{z}\,dz$, which is independent of α. Consequently, $F(\alpha) = 6$.

The difficulty encountered here is essentially of the same sort as that appearing in the differentiation of series term by term. In fact, it will be shown next that a study of the *convergent* improper integrals of the form

$$F(\alpha) = \int_a^\infty f(x, \alpha)\, dx, \tag{96-4}$$

where $f(x, \alpha)$ is a continuous function for $x \geq a$ and for every value of α in the interval $\alpha_0 \leq \alpha \leq \alpha_1$, is equivalent to the study of a series of integrals.

Let

$$a \equiv a_1 < a_2 < a_3 < \cdots < a_n < \cdots$$

be any increasing sequence of numbers such that

$$\lim_{n \to \infty} a_n = \infty.$$

Then the integral (96-4) can be written as an infinite series of integrals, namely,*

$$\begin{aligned} \int_a^\infty f(x, \alpha)\, dx &= \int_{a=a_1}^{a_2} f(x, \alpha)\, dx + \int_{a_2}^{a_3} f(x, \alpha)\, dx + \cdots \\ &\qquad + \int_{a_n}^{a_{n+1}} f(x, \alpha)\, dx + \cdots \\ &= \sum_{n=1}^{\infty} \int_{a_n}^{a_{n+1}} f(x, \alpha)\, dx. \end{aligned} \tag{96-5}$$

Now each term of the series in the right-hand member of (96-5) is a continuous function, so that

$$\int_a^\infty f(x, \alpha)\, dx \equiv \sum_{n=1}^{\infty} u_n(\alpha), \tag{96-6}$$

---

* It is important to note that the requirement that the integral be convergent is essential if one is to reduce the study of the integral to that of the series. For example, $\int_0^\infty \sin x\, dx$ is divergent, but the series

$$\sum_{n=0}^{\infty} u_n = \int_0^{2\pi} \sin x\, dx + \int_{2\pi}^{4\pi} \sin x\, dx + \cdots + \int_{2n\pi}^{2(n+1)\pi} \sin x\, dx + \cdots$$

has its sum equal to zero. Accordingly, $\int_0^\infty \sin x\, dx \neq \sum_{n=0}^{\infty} u_n$.

where

$$u_n(\alpha) \equiv \int_{a_n}^{a_{n+1}} f(x, \alpha)\, dx, \qquad (\alpha_0 \leq \alpha \leq \alpha_1).$$

If the series $\sum_{n=1}^{\infty} u_n(\alpha)$ is uniformly convergent, then the function $F(\alpha)$ will surely be continuous. Accordingly, one can borrow the terminology of infinite series and say that the integral (96-4) will define a continuous function if it is uniformly convergent. This, of course, means that the series (96-6), which represents the integral, is uniformly convergent. Inasmuch as the construction of the series (96-6) is not unique, it is more convenient to define the uniform convergence of the integral (96-4) as follows:

Definition. *The integral* $\int_a^\infty f(x, \alpha)\, dx$ *converges uniformly in the interval* $\alpha_0 \leq \alpha \leq \alpha_1$ *if, corresponding to any* $\epsilon > 0$, *one can find a number* N, *independent of* α, *such that*

$$\left|\int_p^\infty f(x, \alpha)\, dx\right| < \epsilon$$

for every value of α *in the interval* $(\alpha_0, \alpha_1,)$ *whenever* $p \geq N$.

If the integral $\int_a^\infty f(x, \alpha)\, dx$ converges uniformly, then the series (96-6) will also converge uniformly, so that for any $\epsilon > 0$, one can find a number N independent of α such that

$$R_n \equiv \left|\sum_{i=n}^{\infty} u_{i+1}(\alpha)\right| < \epsilon,$$

whenever $n \geq N$.

As in the case of uniform convergence of series, the definition does not provide a useful test for uniform convergence. A simpler test for uniform convergence of integrals follows directly from the Weierstrass test for series.

Weierstrass Test for Integrals. *If* $\varphi(x)$ *is a positive and continuous function for* $x \geq N$, *and if* $|f(x, \alpha)| \leq \varphi(x)$ *for* $\alpha_0 \leq \alpha \leq \alpha_1$ *and* $x \geq N$, *then*

$$\int_N^\infty f(x, \alpha)\, dx$$

is uniformly and absolutely convergent for every α *in the interval* (α_0, α_1) *provided that* $\int_N^\infty \varphi(x)\,dx$ *converges.*

It was remarked earlier that the function $F(\alpha)$, defined by the integral (96-4), will surely be continuous in the closed interval (α_0, α_1) if the integral converges uniformly.

Using the notion of uniform convergence of improper integrals and noting Theorems 2 and 3, Sec. 70, it is easy to establish the following sufficient condition for integrating and differentiating under the integral sign.

Theorem. *Let* $f(x, \alpha)$ *be continuous in* x *and* α *for* $\alpha_0 \leq \alpha \leq \alpha_1$ *and for all values of* $x \geq a$*; then*

$$F(\alpha) \equiv \int_a^\infty f(x, \alpha)\,dx$$

can be integrated under the integral sign with respect to α *in any interval* (β_1, β_2) *contained in* (α_0, α_1) *to yield*

$$\int_{\beta_1}^{\beta_2} F(\alpha)\,d\alpha = \int_a^\infty \int_{\beta_1}^{\beta_2} f(x, \alpha)\,d\alpha\,dx,$$

if $\int_a^\infty f(x, \alpha)\,dx$ *is uniformly convergent in* (α_0, α_1).

Moreover, if $\int_a^\infty \frac{\partial f(x, \alpha)}{\partial \alpha}\,dx$ *converges uniformly for every value of* α *in the interval* (α_0, α_1) *and* f_α *is continuous, then*

$$F'(\alpha) = \int_a^\infty \frac{\partial f(x, \alpha)}{\partial \alpha}\,dx.$$

Some applications of this theorem to the evaluation of improper integrals are indicated in the next section.

PROBLEMS

1. Establish the uniform convergence of the following integrals:

(a) $\int_0^\infty e^{-x^2} \cos \alpha x\,dx;$

(b) $\int_0^\infty e^{-\alpha x^2}\,dx, \qquad \alpha > 0;$

(c) $\int_0^\infty e^{-\alpha x}\, x^3\,dx, \qquad \alpha > 0.$

2. Integrate $\int_a^b \cos x^2\,dx$, $(b > a)$, by parts and obtain

$$\int_a^b \cos x^2\,dx = \frac{1}{2b} \sin b^2 - \frac{1}{2a} \sin a^2 + \frac{1}{2}\int_a^b \frac{\sin x^2}{x^2}\,dx.$$

Deduce from the result the convergence of the integral $\int_0^\infty \cos x^2\,dx$.

3. Use integration by parts to show that $\int_1^\infty \cos x \log x\,dx$ diverges.

4. Use integration by parts to show that $\int_0^1 \cos x \log x\,dx$ converges to

$$-\int_0^1 \frac{\sin x}{x}\,dx.$$

97. Evaluation of Improper Integrals. This section is concerned with the evaluation of several improper integrals that cannot be calculated directly with the aid of the fundamental theorem of the integral calculus. It will be seen that the theorem of the preceding section plays an important part in nearly all the examples considered below.

Example 1. Let it be required to evaluate the probability integral

$$I = \int_0^\infty e^{-x^2}\,dx. \tag{97-1}$$

The convergence of this integral has already been established, in Sec. 93, but it will be profitable to attack the problem in a different way which will lead to an estimate of the magnitude of the number I.

Now

$$\int_0^\infty e^{-x^2}\,dx = \int_0^1 e^{-x^2}\,dx + \int_1^\infty e^{-x^2}\,dx.$$

But, if x lies between 0 and 1, then

$$\int_0^1 xe^{-x^2}\,dx < \int_0^1 e^{-x^2}\,dx,$$

and

$$\int_0^1 xe^{-x^2}\,dx = -\frac{1}{2}e^{-x^2}\Big|_0^1 = \frac{1}{2}\left(1 - \frac{1}{e}\right).$$

Furthermore, if $0 \le x \le 1$, then $e^{-x^2} \le 1$, so that

$$\int_0^1 e^{-x^2}\,dx < 1.$$

Consequently,

$$\frac{1}{2}\left(1 - \frac{1}{e}\right) < \int_0^1 e^{-x^2}\,dx < 1. \tag{97-2}$$

Similarly, if $1 \leq x < \infty$, then

$$0 < \int_1^\infty e^{-x^2}\,dx < \int_1^\infty xe^{-x^2}\,dx.$$

The integral $\int_1^\infty xe^{-x^2}\,dx$ can be calculated with the aid of the fundamental theorem to give

$$\int_1^\infty xe^{-x^2}\,dx = \frac{1}{2e}.$$

Hence,

$$0 < \int_1^\infty e^{-x^2}\,dx < \frac{1}{2e}. \tag{97-3}$$

It follows from (97-2) and (97-3) that

$$\frac{1}{2}\left(1 - \frac{1}{e}\right) < \int_0^\infty e^{-x^2}\,dx < 1 + \frac{1}{2e}.$$

The exact value of (97-1) will be obtained with the aid of the following ingenious device. Since I is not a function of the variable of integration, one can write

$$I = \int_0^\infty e^{-y^2}\,dy. \tag{97-4}$$

Multiplying (97-1) by (97-4) gives

$$\begin{aligned} I^2 &= \int_0^\infty e^{-x^2}\,dx \int_0^\infty e^{-y^2}\,dy \\ &= \int_0^\infty dx \int_0^\infty e^{-(x^2+y^2)}\,dy \\ &= \int_0^\infty \int_0^\infty e^{-(x^2+y^2)}\,dx\,dy. \end{aligned}$$

Transforming to polar coordinates by means of

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta, \end{aligned}$$

and noting that the integration extends over the area of the first quadrant of the xy-plane gives*

$$I^2 = \int_0^{\frac{\pi}{2}} \int_0^\infty e^{-r^2} r\,dr\,d\theta = \frac{\pi}{4}.$$

---

* See Sec. 46.

Hence,

$$I = \frac{\sqrt{\pi}}{2}.$$

The same device can be used to evaluate

$$\int_0^\infty x^2 e^{-x^2}\, dx.$$

Example 2. Evaluate

$$\varphi(\alpha) = \int_0^\infty x^3 e^{-\alpha x^2}\, dx$$

for $\alpha \geq \alpha_0 > 0$. Since for any $\alpha > 0$ there exists a number N such that $e^{-\alpha x^2} \leq \frac{1}{x^5}$ for $x \geq N$, it follows that $x^3 e^{-\alpha x^2} \leq \frac{1}{x^2}$ for $x \geq N$. Moreover,

$$\int_N^\infty \frac{dx}{x^2}$$

converges, and therefore, the given integral converges uniformly. Then, assuming the existence of the integral,

$$\int_\alpha^\infty \varphi(\alpha)\, d\alpha = \int_0^\infty \int_\alpha^\infty x^3 e^{-\alpha x^2}\, d\alpha\, dx = \int_0^\infty x e^{-\alpha x^2}\, dx = \frac{1}{2\alpha}.$$

Hence,

$$\varphi(\alpha) = \frac{1}{2\alpha^2}.$$

Example 3. Evaluate

$$\varphi(\alpha) = \int_0^\infty e^{-x^2} \cos \alpha x\, dx.$$

Now

$$\varphi'(\alpha) = \int_0^\infty -\, x e^{-x^2} \sin \alpha x\, dx$$

and, since

$$|x e^{-x^2} \sin \alpha x| \leq x e^{-x^2},$$

whose integral is convergent, the differentiation is justified.

By integration by parts,

$$\varphi'(\alpha) = \frac{1}{2} e^{-x^2} \sin \alpha x \Big|_0^\infty - \frac{\alpha}{2} \int_0^\infty e^{-x^2} \cos \alpha x\, dx = -\frac{\alpha}{2} \varphi(\alpha).$$

But if

$$\frac{d\varphi}{d\alpha} = -\frac{\alpha\varphi}{2},$$

then

$$\frac{d\varphi}{\varphi} = -\frac{\alpha}{2}\,d\alpha$$

or

$$\log\varphi = -\frac{\alpha^2}{4} + C_1$$

and

$$\varphi = Ce^{-\frac{\alpha^2}{4}}.$$

In order to determine C, consider

$$\varphi(0) = \int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}. \qquad \text{(See Example 1.)}$$

But

$$\varphi(0) = Ce^0 = C,$$

and, therefore,

$$\varphi(\alpha) = \frac{\sqrt{\pi}}{2}e^{-\frac{\alpha^2}{4}}.$$

Example 4. Evaluate

$$F(\alpha, \beta) = \int_0^\infty e^{-\alpha x}\frac{\sin\beta x}{x}\,dx, \qquad \text{where} \qquad \alpha \geq 0. \tag{97-5}$$

Without loss of generality β may be assumed to be positive, since a change in the sign of β merely changes the sign of the integral. Now

$$\frac{\partial F(\alpha, \beta)}{\partial\beta} = \int_0^\infty e^{-\alpha x}\cos\beta x\,dx, \tag{97-6}$$

and, since for a fixed $\alpha > 0$,

$$|e^{-\alpha x}\cos\beta x| \leq e^{-\alpha x},$$

which is independent of β, the differentiation under the sign is valid.

The evaluation of (97-6) is easy. For

$$\frac{\partial F}{\partial \beta} = \int_0^\infty e^{-\alpha x} \cos \beta x \, dx = e^{-\alpha x} \frac{\beta \sin \beta x - \alpha \cos \beta x}{\alpha^2 + \beta^2}\Bigg|_0^\infty$$
$$= \frac{\alpha}{\alpha^2 + \beta^2}.$$

Hence,

$$F(\alpha, \beta) = \int \frac{\alpha}{\alpha^2 + \beta^2} \, d\beta = \tan^{-1} \frac{\beta}{\alpha} + C.$$

But $F(\alpha, 0) = 0$ and, since (97-5) defines a continuous function of β,

$$\lim_{\beta \to 0} F(\alpha, \beta) = 0 = \tan^{-1} \frac{0}{\alpha} + C.$$

Thus

$$C = 0.$$

Consequently,

(97-7) $$F(\alpha, \beta) = \tan^{-1} \frac{\beta}{\alpha}, \qquad \alpha > 0.$$

It follows from (97-5) that

(97-8) $$F(0, \beta) = \int_0^\infty \frac{\sin \beta x}{x} \, dx,$$

and if it is demonstrated that (97-5) defines a continuous function of α (for all $\alpha \geq 0$ and for a fixed β), then (97-7) will afford a means of calculating the integral (97-8). But, by breaking the interval of integration in (97-5) into subintervals of lengths $\frac{\pi}{\beta}$, (97-5) can be written as

(97-9) $$F(\alpha, \beta) = \sum_{n=0}^{\infty} \int_{\frac{n\pi}{\beta}}^{\frac{(n+1)\pi}{\beta}} e^{-\alpha x} \frac{\sin \beta x}{x} \, dx,$$

and it is not difficult to show* that the numerical value of the

---

* If the substitution $x = z + \frac{n\pi}{\beta}$ is made in the integral, the assertion follows almost immediately.

nth term of the series is less than $\frac{1}{n}$. Furthermore, the series is alternating and decreasing, so that the sum of the terms beginning with the $(n + 1)$st is less than $\frac{1}{n+1}$, which is independent of α. Hence, the series (97-9) converges uniformly and thus defines a continuous function for $\alpha \geq 0$.* Accordingly,

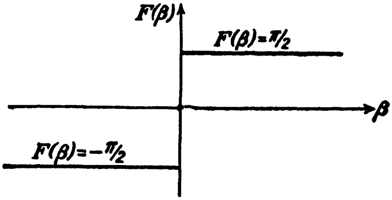

FIG. 77.

$$
\begin{aligned}
F(0, \beta) \equiv \int_0^\infty \frac{\sin \beta x}{x}\, dx &= \lim_{\alpha \to 0+} \tan^{-1} \frac{\beta}{\alpha} \\
&= \frac{\pi}{2}, \qquad \text{if} \quad \beta > 0, \\
&= -\frac{\pi}{2}, \qquad \text{if} \quad \beta < 0, \\
&= 0, \qquad \text{if} \quad \beta = 0.
\end{aligned}
$$

The function defined by $\int_0^\infty \frac{\sin \beta x}{x}\, dx$ is discontinuous at $\beta = 0$ (see Fig. 77). This integral is of importance in the theory of Fourier series.

Example 5. As an example of an integral with a discontinuous integrand, consider

$$
(97\text{-}10) \qquad I = \int_0^1 \frac{\log x}{1 - x}\, dx.
$$

---

* It should be noted that the continuity of $F(\alpha, \beta)$ (for a fixed β) does not ensure the uniform convergence of the integral (97-5). As a matter of fact, the integral is uniformly convergent if $\alpha \geq 0$, but this has not been demonstrated above. Instead of constructing the series, one could prove the uniform convergence of (97-5) and thus deduce the continuity of $F(\alpha, \beta)$. This can be done with the aid of the second mean-value theorem for integrals. See, for example, H. S. Carslaw, Fourier Series and Integrals, 2d ed., p. 184.

Now

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \frac{x^{n+1}}{1-x};$$

hence,

$$\int_0^1 \frac{\log x}{1-x}\,dx = \int_0^1 \log x\,dx + \int_0^1 x \log x\,dx + \cdots + \int_0^1 x^n \log x\,dx + \int_0^1 \frac{x^{n+1}}{1-x} \log x\,dx.$$

If it is established that

$$\lim_{n\to\infty} R_n \equiv \lim_{n\to\infty} \int_0^1 \frac{x^{n+1}}{1-x} \log x\,dx = 0,$$

then

$$\text{(97-11)} \qquad \int_0^1 \frac{\log x}{1-x}\,dx = \sum_{n=0}^{\infty} \int_0^1 x^n \log x\,dx = -\sum_{n=1}^{\infty} \frac{1}{n^2},$$

where the last step results from integration by parts.

The proof that $\lim_{n\to\infty} R_n = 0$ is accomplished without much trouble if one notes that

$$\text{(97-12)} \qquad \int_0^1 \frac{x^{n+1}}{1-x} \log x\,dx = \int_0^a \frac{x^{n+1}}{1-x} \log x\,dx + \int_a^1 \frac{x^{n+1}}{1-x} \log x\,dx,$$

where a is a positive number less than unity. Now

$$\int_0^a \frac{x^{n+1}}{1-x} \log x\,dx < a^{n+1} \int_0^a \frac{\log x}{1-x}\,dx,$$

and, since $a < 1$, the expression* on the right tends to zero as $n \to \infty$.

---

* Note Prob. 1 (j), Sec. 95.

The reader will show that $\frac{\log x}{1-x}$ is bounded if $0 < a \leq x \leq 1$, and hence, the second integral in the right-hand member of (97-12) likewise tends to zero as $n \to \infty$. This completes the proof of the validity of formula (97-11).

The function

$$\zeta(x) \equiv \sum_{n=1}^{\infty} \frac{1}{n^x}$$

has been studied by Riemann and is called the *Riemann zeta-function.*[*] Thus the value of I is $-\zeta(2)$.

PROBLEMS

1. Verify the values of the integrals given below:

(a) $\int_0^\infty e^{-\alpha x^2}\,dx = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}, \qquad \alpha > 0;$

(b) $\int_0^1 x^\alpha\,dx = \frac{1}{\alpha+1}, \qquad \alpha > -1;$

(c) $\int_0^\infty e^{-\left(x-\frac{\alpha}{x}\right)^2}\,dx = \frac{\sqrt{\pi}}{2},$

(d) $\int_0^\infty x^2 e^{-x^2}\,dx = \frac{\sqrt{\pi}}{4};$

(e) $\int_0^\infty e^{-\alpha x}\sin x\,dx = \frac{1}{1+\alpha^2}, \qquad \alpha > 0;$

(f) $\int_0^\infty \frac{e^{-\alpha x} - e^{-\beta x}}{x}\,dx = \log\frac{\beta}{\alpha}, \ \alpha > 0, \beta > 0;$

(g) $\int_0^\infty e^{-\alpha^2 x^2}\cos\beta x\,dx = \frac{\sqrt{\pi}}{2\alpha}e^{-\beta^2/4\alpha^2}, \qquad \alpha > 0;$

(h) $\int_0^\infty x e^{-\alpha x^2}\,dx = \frac{1}{2\alpha}, \qquad \alpha > 0.$

2. Let

$$F(\alpha) = \int_0^\alpha \frac{\log(1+\alpha x)}{1+x^2}\,dx.$$

---

[*] See, for example, E. T. Whittaker and G. N. Watson, A Course of Modern Analysis, Chap. XIII.

Differentiate this integral with respect to α, and evaluate the resulting integral. Hence, show that

$$\int_0^1 \frac{\log(1+x)}{1+x^2}\,dx = \frac{\pi \log 2}{8}.$$

3. Assuming that

$$\int_0^\infty e^{-\alpha x^2}\,dx = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}, \qquad \alpha > 0,$$

is valid for complex values of $\alpha \equiv a + bi = re^{i\theta}$, where $a \geq 0$, show that

$$\int_0^\infty e^{-(a+bi)x^2}\,dx = \frac{1}{2}\sqrt{\frac{\pi}{r}}\,e^{-i(\theta/2)}.$$

Set $a = 0$ and $b = 1$, and obtain the Fresnel integrals

$$\int_0^\infty \cos x^2\,dx = \int_0^\infty \sin x^2\,dx = \frac{\sqrt{\pi}}{2\sqrt{2}}.$$

4. Using the formula

$$\frac{1}{\alpha^3} = \frac{1}{2}\int_0^\infty x^2 e^{-\alpha x}\,dx,$$

show that

$$\zeta(3) \equiv \sum_{n=1}^{\infty} \frac{1}{n^3} = \frac{1}{2}\int_0^\infty \frac{x^2\,dx}{e^x - 1}.$$

5. Show that

$$\int_0^\infty \frac{\sin^2 x}{x^2}\,dx = \frac{\pi}{2}.$$

Hint: Integrate by parts, and use $\displaystyle\int_0^\infty \frac{\sin \beta x}{x}\,dx = \frac{\pi}{2}$, if $\beta > 0$.

98. Improper Multiple Integrals. Let $f(x, y)$ be a function of two variables x and y which is continuous at every point of a two-dimensional region R, bounded by a closed curve C, except at the point $Q(\xi, \eta)$ where $f(x, y)$ becomes infinite. The symbol

$$\int_R f(x, y)\,dS, \tag{98-1}$$

where dS is the element of area of the region R, is called an *improper double integral.*

Let an arbitrary infinite sequence of closed curves C_1, C_2, $\cdots$, C_n, $\cdots$ be drawn in R in such a way that each curve C_n encloses the point $Q(\xi, \eta)$ and where each curve lies entirely within the preceding one. Furthermore, let the greatest linear dimension of the region R_n enclosed by C_n tend to zero as n increases indefinitely (Fig. 78). Denote the closed region which is exterior to C_n and interior to C by

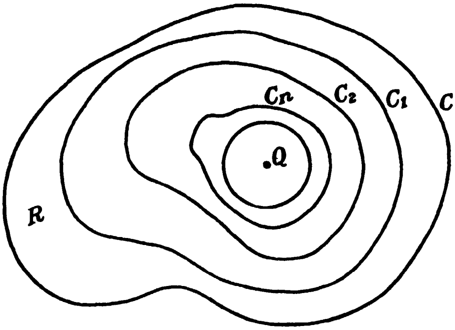

FIG. 78.

$$R - R_n, \qquad (n = 1, 2, \cdots);$$

then the function $f(x, y)$ will be continuous in the regions $R - R_n$, and the double integral

$$\int_{R-R_n} f(x, y)\, dS$$

will exist for every value of n.

If

$$\lim_{n\to\infty} \int_{R-R_n} f(x, y)\, dS$$

exists for an arbitrary choice of the regions R_n, then (98-1) is said to converge. It follows from the definition of the limit that the convergence of (98-1) means that for any $\epsilon > 0$ one can find a positive number δ such that

$$\text{(98-2)} \qquad \left| \int_{R'} f(x, y)\, dS \right| < \epsilon,$$

for every subregion R' of R which does not contain Q and which lies within the circle of radius δ with center at Q.

If $f(x, y)$ contains a parameter α, where $\alpha_0 \leq \alpha \leq \alpha_1$, then the convergent improper integral defines a function of α. Moreover, if the choice of the radius δ of the circle is independent of the choice of α in (α_0, α_1), then the integral

$$\int_R f(x, y, \alpha)\, dS$$

will be said to converge uniformly. A uniformly convergent double integral defines a continuous function of α, and it may be integrated under the integral sign with respect to α. The differentiation under the integral sign is permissible if the derived integral, namely,

$$\int_R \frac{\partial f}{\partial \alpha}\, dS$$

converges uniformly.

These considerations can be extended to cover the case of improper double integrals containing a finite number of parameters. An extension of the definition of an improper integral to functions of more than two variables is likewise immediate.

Thus, if $f(x, y, z)$ becomes infinite at the point $Q(\xi, \eta, \zeta)$, but is continuous at every other point of a three-dimensional region R containing Q, then the symbol

$$\int_R f(x, y, z)\, dV, \tag{98-3}$$

where dV is the element of volume, is called an *improper triple integral.* If R_n (where $n = 1, 2, \cdots$) is a sequence of regions bounded by the closed surfaces S_n, which are constructed in a manner entirely analogous to the construction of the curves C_n in the definition of the two-dimensional case, then (98-3) is said to be a convergent triple integral provided that

$$\lim_{n \to \infty} \int_{R - R_n} f(x, y, z)\, dV$$

exists independently of the choice of the regions R_n.

The following test, giving a sufficient condition for the convergence of an improper triple integral, is used frequently in applied mathematics.

Test for Convergence. *Let $f(x, y, z)$ be a continuous function at every point $P(x, y, z)$ of the region R with the exception of the* *point* $Q(\xi, \eta, \zeta)$, *where* $f(x, y, z)$ *becomes infinite. If there exists a positive constant* A *such that*

$$|r^n f(x, y, z)| < A, \tag{98-4}$$

where $r = \sqrt{(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2}$ *and* $n < 3$, *then the integral*

$$\int_R f(x, y, z)\, dV \tag{98-5}$$

is convergent.

The proof of this test follows from the criterion of convergence analogous to (98-2). Indeed, if the integral (98-5) is to converge, then for an arbitrary $\epsilon > 0$,

$$\left|\int_{R'} f(x, y, z)\, dV\right| < \epsilon$$

for every three-dimensional subregion R' (of R) not containing $Q(\xi, \eta, \zeta)$ and lying within the sphere S of radius δ whose center is at Q. But if the inequality (98-4) is satisfied, then

$$\left|\int_{R'} f(x, y, z)\, dV\right| < \int_{R'} \frac{A}{r^n}\, dV. \tag{98-6}$$

Since R' is only a part of the region bounded by the sphere S, and since the integrand is positive, it follows that

$$\int_{R'} \frac{A}{r^n}\, dV < \int_S \frac{A}{r^n}\, dV. \tag{98-7}$$

The integral on the right of this inequality can be evaluated easily in spherical coordinates, since*

$$\begin{aligned}\int_S \frac{A}{r^n}\, dV &= A\int_0^{2\pi}\int_0^{\pi}\int_0^{\delta} r^{2-n} \sin\theta\, dr\, d\theta\, d\varphi \\ &= A\frac{4\pi\delta^{3-n}}{3 - n}.\end{aligned}$$

But if $n < 3$, then the magnitude of the right-hand member of (98-7) can be made as small as desired by choosing the radius δ small enough. It follows from (98-7) and (98-6) that

$$\left|\int_{R'} f(x, y, z)\, dV\right|$$

---

* See Sec. 49.

can be made as small as desired, and this establishes the convergence of (98-5).

The reader will show, by using a similar argument, that an improper double integral will surely converge if

$$|f(x, y)| < \frac{A}{r^n},$$

where $n < 2$, A is a positive constant, and r is the distance from P to Q.

As an illustration of the application of the theory of this section, let it be required to calculate the gravitational attraction on a unit mass located at a point $P(x, y, z)$ due to a distribution of mass of density ρ contained within the volume V (Fig. 79).

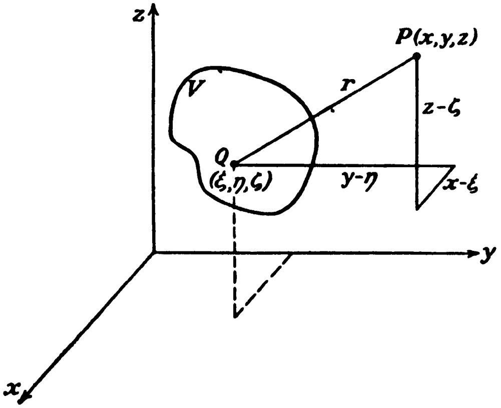

Fig. 79.

Denote the coordinates of the points of the volume V by (ξ, η, ζ); then the density ρ is a continuous function of ξ, η, and ζ and the element of mass dm of the body is

$$dm = \rho\, dV.$$

The attraction at the point P, due to an element of mass dm, is given by the inverse square law, namely,

$$\frac{dm}{r^2} = \frac{\rho(\xi, \eta, \zeta)}{r^2}\, dV,$$

where $r = \sqrt{(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2}$. The components of the force of attraction along the coordinate axes are*

$$-\frac{dm}{r^2}\cdot\frac{x-\xi}{r}, \qquad -\frac{dm}{r^2}\cdot\frac{y-\eta}{r}, \qquad -\frac{dm}{r^2}\cdot\frac{z-\zeta}{r};$$

hence the force X in the direction of the x-axis is

$$X = \int_V \frac{\rho(\xi, \eta, \zeta)(\xi - x)}{r^3}\, dV, \tag{98-8}$$

with similar expressions for the components of force in the directions of the y- and z-axes.

The variables of integration in (98-8) are ξ, η, and ζ, so that the element of volume dV can be expressed as $d\xi\, d\eta\, d\zeta$. If the point P is outside the volume V, so that r never vanishes, then the integral (98-8) is proper. If, however, P is chosen within the body, then the denominator of the integrand in (98-8) vanishes when $\xi = x$, $\eta = y$, and $\zeta = z$, but it is not difficult to show that the resulting improper triple integral converges. Indeed, let the maximum value of $\rho(\xi, \eta, \zeta)$ be A; then

$$\left|\frac{\rho}{r^2}\cdot\frac{\xi - x}{r}\right| \le \frac{A}{r^2}$$

and, since the exponent of r is less than 3, the integral (98-8) will surely converge.

The potential $\Phi(x, y, z)$ at the point P due to a distribution of mass of density ρ is defined by the equation†

$$\Phi(x, y, z) = \int_V \frac{\rho(\xi, \eta, \zeta)}{r}\, dV, \tag{98-9}$$

and (98-9) is certainly convergent at all interior points if ρ is a continuous function.

It is important to note that the maximum density A is independent of the location of the point $P(x, y, z)$, so that the integrals (98-8) and (98-9) converge uniformly for all values of the parameters x, y, and z. Accordingly, the potential Φ and the components of force X, Y, and Z are continuous functions of x, y, and z throughout all space.

---

* See p. 202.

† See p. 203

Since

$$\frac{\partial r}{\partial x} = \frac{x - \xi}{r}, \qquad \frac{\partial r}{\partial y} = \frac{y - \eta}{r}, \qquad \frac{\partial r}{\partial z} = \frac{z - \zeta}{r},$$

it follows from (98-8) that

$$(98\text{-}10) \qquad \begin{cases} X = \displaystyle\int_V \rho \frac{\partial\left(\frac{1}{r}\right)}{\partial x}\, dV, \\ Y = \displaystyle\int_V \rho \frac{\partial\left(\frac{1}{r}\right)}{\partial y}\, dV, \\ Z = \displaystyle\int_V \rho \frac{\partial\left(\frac{1}{r}\right)}{\partial z}\, dV, \end{cases}$$

which are precisely the expressions that one obtains by differentiating (98-9), under the integral sign, with respect to the parameters x, y, and z. The differentiation under the integral sign is clearly justified since (98-8) converges uniformly. Consequently, one can write*

$$(98\text{-}11) \qquad \frac{\partial \Phi}{\partial x} = X, \qquad \frac{\partial \Phi}{\partial y} = Y, \qquad \frac{\partial \Phi}{\partial z} = Z.$$

If the point P is outside the volume V, the integrals (98-10) are proper and hence can be differentiated under the sign. Thus,

$$\frac{\partial X}{\partial x} = \int \rho \frac{\partial^2\left(\frac{1}{r}\right)}{\partial x^2}\, dV,$$

with similar expressions for $\frac{\partial Y}{\partial y}$ and $\frac{\partial Z}{\partial z}$. But

$$\frac{\partial^2\left(\frac{1}{r}\right)}{\partial x^2} = \frac{3(x - \xi)^2}{r^5} - \frac{1}{r^3},$$

---

* See p. 203.

and the expressions for $\dfrac{\partial^2\left(\dfrac{1}{r}\right)}{\partial y^2}$ and $\dfrac{\partial^2\left(\dfrac{1}{r}\right)}{\partial z^2}$ are similar. Hence.

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = 0,$$

and a reference to (98-11) shows that

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0. \tag{98-12}$$

This is the celebrated partial differential equation of Laplace.

If the point P is interior to the body, the differentiation of (98-10) under the integral sign is not legitimate, and the equation (98-12) is no longer satisfied unless $\rho = 0$.

PROBLEMS

1. Prove that the force of attraction at a point P, exterior to a homogeneous sphere, is $F = \dfrac{M}{R^2}$, where R is the distance from the center of the sphere to the point P and M is the mass of the sphere.

2. If the force of attraction exerted by an element of mass dm on a point is $\dfrac{k\,dm}{r^2}$, where r is the distance from the element of mass to the point, find the attraction of (*a*) a homogeneous right-circular cone upon a point at the vertex; (*b*) a homogeneous right-circular cylinder upon a point on its axis.

Ans. (*a*) $2\pi k\rho h(1 - \cos\alpha)$, where h is the altitude and 2α is the angle at the vertex;

(*b*) $2\pi k\rho[h + \sqrt{R^2 + a^2} - \sqrt{(R + h)^2 + a^2}\,]$, where h is the altitude and a is the radius of the cylinder, and R is the distance from the point to one base of the cylinder.

3. Find the force of attraction on a unit mass located at a point within the cavity of a homogeneous spherical shell.

4. Find the potential due to a homogeneous spherical shell at all points exterior to the shell, and within the cavity.

5. Find the potential at the points on the axis of a homogeneous cylindrical shell.

99. Gamma Functions. An interesting function, which provides a generalization of the factorial, is defined by the improper integral

$$(99\text{-}1) \qquad \Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx, \qquad \text{where} \qquad \alpha > 0,$$

and is called the *Gamma function.*

If α is a positive number less than unity, the integrand of (99-1) becomes infinite at the lower limit $x = 0$, so that it becomes necessary to investigate the behavior of the integral at both limits. Accordingly, (99-1) will be written as the sum of two integrals, namely,

$$(99\text{-}2) \qquad \Gamma(\alpha) = \int_0^1 x^{\alpha-1}e^{-x}\,dx + \int_1^\infty x^{\alpha-1}e^{-x}\,dx,$$

the second of which has a continuous integrand for all values of $x \geq 1$. Moreover, this second integral converges if $\alpha > 0$ since*

$$\lim_{x\to\infty} x^2(e^{-x}x^{\alpha-1}) = \lim_{x\to\infty} x^{\alpha+1}e^{-x} = 0.$$

If $\alpha \geq 1$, the first integral is proper, since its integrand is continuous. It remains to investigate the behavior of the integral

$$(99\text{-}3) \qquad \int_0^1 x^{\alpha-1}e^{-x}\,dx$$

when $0 < \alpha < 1$. Now

$$\lim_{x\to 0} x^k(x^{\alpha-1}e^{-x}) = 0, \qquad \text{if} \qquad k + \alpha - 1 > 0,$$

and this condition is surely satisfied if $1 - \alpha < k < 1$. It follows from the corollary to the theorem of Sec. 95 that (99-3) is a convergent integral.

It is also clear that (99-1) diverges if $\alpha \leq 0$, so that (99-1) defines a function for positive values of α only. However, it is possible to define the function $\Gamma(\alpha)$ for negative values of α with the aid of the recursion formula which will be developed next.

If $\alpha > 0$, then it follows from (99-1) that

$$(99\text{-}4) \qquad \Gamma(\alpha + 1) = \int_0^\infty x^\alpha e^{-x}\,dx.$$

---

* See corollary to Theorem 1, Sec. 93.

Integrating the right-hand member of (99-4) by parts gives

$$\begin{aligned}\int_0^\infty x^\alpha e^{-x}\,dx &= -x^\alpha e^{-x}\Big|_0^\infty + \alpha\int_0^\infty x^{\alpha-1}e^{-x}\,dx \\ &= \alpha\int_0^\infty x^{\alpha-1}e^{-x}\,dx \\ &= \alpha\Gamma(\alpha).\end{aligned}$$

Thus,

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha). \tag{99-5}$$

But

$$\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1,$$

so that when $\alpha = 1$, the formula (99-5) becomes

$$\Gamma(2) = 1\cdot\Gamma(1) = 1.$$

Setting $\alpha = 2, 3, \cdots, n$ gives

$$\begin{aligned}\Gamma(3) &= 2\Gamma(2) = 1\cdot 2, \\ \Gamma(4) &= 3\Gamma(3) = 1\cdot 2\cdot 3, \\ &\cdots\cdots\cdots\cdots\cdots, \\ \Gamma(n) &= (n-1)\Gamma(n-1) = (n-1)!, \\ \Gamma(n+1) &= n\Gamma(n) = n!.\end{aligned}$$

Hence, the formula (99-5) enables one to compute the values of $\Gamma(\alpha)$ for all positive integral values of the argument α.

If by some means (for example, by using infinite series) the values of $\Gamma(\alpha)$ are obtained for all values of α between 1 and 2, then, with the aid of the recursion formula (99-5), the values of $\Gamma(\alpha)$ are readily obtained when α lies between 2 and 3. Knowing these values, it is easy to obtain $\Gamma(\alpha)$ where $3 < \alpha < 4$, etc. The values of $\Gamma(\alpha)$ for α lying between 1 and 2 have been computed* to a high degree of accuracy, so that it is possible to find the value of $\Gamma(\alpha)$ for all $\alpha > 0$.

It remains to define $\Gamma(\alpha)$ for negative values of α. The recursion formula (99-5) can be written as

$$\Gamma(\alpha) = \frac{\Gamma(\alpha+1)}{\alpha}. \tag{99-6}$$

The formula (99-6) becomes meaningless when α is set equal to zero, since

$$\lim_{\alpha\to 0+}\Gamma(\alpha) = +\infty \qquad\text{and}\qquad \lim_{\alpha\to 0-}\Gamma(\alpha) = -\infty.$$

---

* Small table is found in Peirce's "Table of Integrals," p. 140.

It follows from (99-6) that the function $\Gamma(-\alpha)$ is discontinuous when α is a positive integer.

If any number $-1 < \alpha < 0$ is substituted in the left-hand side of (99-6), the right-hand side gives the value of $\Gamma(-\alpha)$, since the values of $\alpha + 1$ lie between 0 and 1, and $\Gamma(\alpha)$ is known for these values of α. Thus,

$$\Gamma\left(-\frac{1}{2}\right) = \frac{\Gamma(\frac{1}{2})}{-\frac{1}{2}}, \qquad \Gamma(-0.9) = \frac{\Gamma(0.1)}{-0.9}, \text{ etc.}$$

In this manner the values of $\Gamma(\alpha)$ for $-1 < \alpha < 0$ can be computed. Knowing these values and using the recursion formula (99-6), the values of $\Gamma(\alpha)$ for $-2 < \alpha < -1$ can be obtained, etc. The adjoining figure represents the graph of $\Gamma(\alpha)$ (Fig. 80).

It was observed that

$$\Gamma(\alpha + 1) = \alpha!$$

when α is a positive integer. This formula may serve as the definition of factorials of fractional numbers. Thus,

$$\Gamma(\tfrac{3}{2}) = \tfrac{1}{2}!,$$
$$\Gamma(\tfrac{1}{2}) = (-\tfrac{1}{2})!,$$
$$\Gamma(1) = 0! = 1.$$

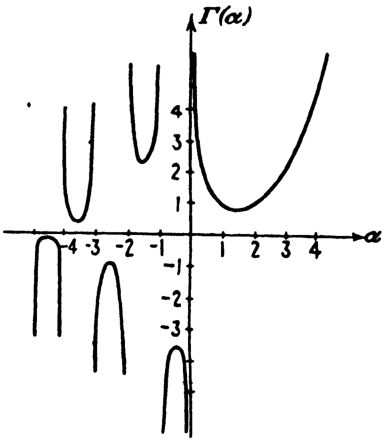

FIG. 80.

This section will be concluded with an ingenious method* of evaluating $\tfrac{1}{2}!$. Now

$$\tfrac{1}{2}! = \Gamma(\tfrac{3}{2}) = \int_0^\infty e^{-x} x^{\frac{1}{2}}\, dx.$$

If the variable in this integral be changed by the transformation $x = y^2$, the integral becomes

$$\tfrac{1}{2}! = 2\int_0^\infty e^{-y^2} y^2\, dy \tag{99-7}$$

Since the definite integral is independent of the variable of integration and is a function of the limits,

---

* See also Sec. 97.

$$\tfrac{1}{2}! = 2\int_0^\infty e^{-z^2}z^2\,dz. \tag{99-8}$$

Multiplying (99-7) by (99-8) gives

$$(\tfrac{1}{2}!)^2 = 4\int_0^\infty e^{-z^2}z^2\,dz\int_0^\infty e^{-y^2}y^2\,dy,$$

which can be written as a double integral

$$(\tfrac{1}{2}!)^2 = 4\int_0^\infty\int_0^\infty e^{-(y^2+z^2)}y^2z^2\,dy\,dz. \tag{99-9}$$

In order to evaluate (99-9), transform it into polar coordinates by setting $z = r\cos\theta$ and $y = r\sin\theta$. The element of area $dy\,dz$ becomes $r\,dr\,d\theta$, and (99-9) becomes

$$(\tfrac{1}{2}!)^2 = 4\int_0^\infty dr\int_0^{\frac{\pi}{2}} r^5e^{-r^2}\sin^2\theta\cos^2\theta\,d\theta.$$

But

$$\int_0^{\frac{\pi}{2}} \sin^2\theta\cos^2\theta\,d\theta = \frac{\pi}{16}$$

and

$$\int_0^\infty e^{-r^2}r^5\,dr = 1.$$

The latter integral is evaluated by integration by parts. Therefore

$$\left(\frac{1}{2}!\right)^2 = \frac{\pi}{4} \quad\text{or}\quad \frac{1}{2}! = \frac{\sqrt{\pi}}{2}.$$

It can be shown with the aid of the recursion formula that

$$(-\tfrac{1}{2})! = \sqrt{\pi}.$$

PROBLEMS

1. Compute the values of $\Gamma(\alpha)$ for every integer and half integer from 0 to 5 by using the relations $\Gamma(1) = 1$ and $\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$. Plot the curve $y = \Gamma(\alpha)$ with the aid of these values.

2. The beta-function $B(m, n)$ is defined by the integral

$$B(m, n) = \int_0^1 x^{m-1}(1 - x)^{n-1}\,dx.$$

If x is replaced by y^2 in $\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}\,dx$, there results

$$\Gamma(n) = 2\int_0^\infty e^{-y^2}y^{2n-1}\,dy.$$

Using this integral, form

$$\Gamma(m)\Gamma(n) = 4\int_0^\infty x^{2n-1}e^{-x^2}\,dx\int_0^\infty y^{2m-1}e^{-y^2}\,dy.$$

Express this product as a double integral, transform to polar coordinates, and show that

$$B(m, n) = B(n, m) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

3. Show, by a suitable change of variable, that (99-1) reduces to

$$\int_0^1 \left(\log\frac{1}{y}\right)^{\alpha-1} dy.$$

4. Show that

$$\frac{d^n\Gamma(\alpha)}{d\alpha^n} = \int_0^\infty x^{\alpha-1}e^{-x}(\log \alpha)^n\,dx,$$

and justify the differentiation of (99-1) under the integral sign.

CHAPTER XI

FOURIER SERIES

100. Criterion of Approximation. It is already known[*] that a suitably restricted function $f(x)$ can be represented in a series of ascending powers of x. The restrictions imposed upon $f(x)$ are quite severe inasmuch as it is assumed that the function not only possesses derivatives of all orders, but also that the remainder in the expansion converges to zero uniformly.[†] It is obvious that a function that is capable of a power-series representation can be approximated arbitrarily closely in a given interval by a polynomial of sufficiently high degree. The criterion of the closeness of approximation in this case consists of the requirement that the numerical value of the deviation of the values of the function from those of the polynomial (that is, the remainder $R_n(x) \equiv f(x) - \sum_{k=0}^{n} a_k x^k$) can be made less than some specified amount for all values of x in the given interval.

The question naturally arises whether it is possible to represent a given function $f(x)$ by a trigonometric polynomial of the form[‡]

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \tag{100-1}$$

where a_k and b_k are real constants so selected as to satisfy some criterion of the goodness of approximation.

It should be observed that since every term in the right-hand member of (100-1) is a periodic function of period 2π, $S_n(x)$

---

[*] See Chap. IX.

[†] See Sec. 82.

[‡] The reason for writing the constant term as $\frac{a_0}{2}$ will appear in the next section.

is also periodic.* Consequently, only periodic functions $f(x)$ of period 2π can be approximated by a polynomial of the form (100-1). Or, what amounts to the same thing, one can restrict the problem of approximating a nonperiodic function $f(x)$ to some interval of width 2π and define the function $f(x)$ outside this interval so that it is periodic. It will be assumed from this point on that the problem of approximating $f(x)$ by a trigonometric polynomial is confined to the interval $(-\pi, \pi)$ and that outside this interval the function $f(x)$ is defined by the equation $f(x + 2\pi) = f(x)$. Of course, any interval $(a, a + 2\pi)$ will do equally well.

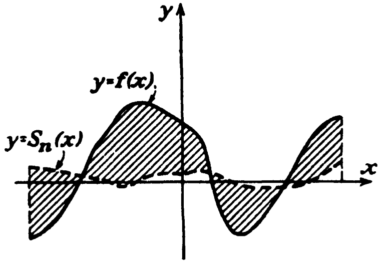

FIG. 81.

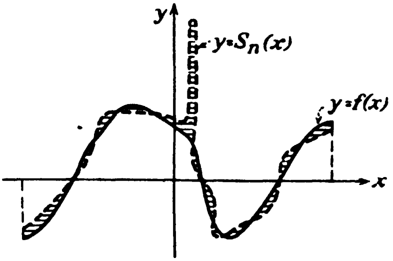

FIG. 82.

In formulating the problem of approximating $f(x)$ by a trigonometric polynomial $S_n(x)$, one could demand that the coefficients a_k and b_k be so chosen as to satisfy the requirement that the difference $\delta_n(x)$, between the values of the function $f(x)$ and the polynomial $S_n(x)$, does not exceed a prescribed amount for all values of x in the interval $(-\pi, \pi)$. However, such a criterion would regard the dotted curve shown in Fig. 81 as constituting a better representation of the function $f(x)$, shown by the solid line, than the wavy dotted line approximating the same function $f(x)$ in Fig. 82. For it is clear that the maximum deviation of the approximating function shown in Fig. 81 is less than the corresponding maximum deviation shown in Fig. 82.

It appears more reasonable to set up as a criterion for the goodness of approximation the magnitude of the integral of the absolute value of $\delta_n(x)$, namely,

$$\int_{-\pi}^{\pi} |\delta_n(x)|\, dx,$$

---

* A function is said to be periodic of period a if $f(x) = f(x + a)$.

since this integral represents the shaded area between the curves $S_n(x)$ and $f(x)$. The closer the approximating curve lies to $f(x)$, the smaller the magnitude of the integral will be. However, a study of integrals containing absolute values of functions in the integrand is quite difficult, and it is customary to consider the integral of the square of $\delta_n(x)$ instead of the integral of its absolute value. Accordingly, it will be agreed that the "best" approximation for the function $f(x)$ by means of a trigonometric polynomial $S_n(x)$ is that polynomial in which the coefficients a_k and b_k are so selected as to make the magnitude of the integral

$$(100\text{-}2)\quad \begin{aligned} I_n &= \int_{-\pi}^{\pi} [\delta_n(x)]^2\,dx \\ &= \int_{-\pi}^{\pi} \left[f(x) - \frac{a_0}{2} - \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \right]^2 dx \end{aligned}$$

a minimum.

Therefore, the problem of determining the coefficients in (100-1) is reduced to that of minimizing the function I_n of $2n + 1$ variables a_k and b_k. The following section is concerned with this problem of minimizing the integral (100-2).

101. Fourier Coefficients. Let $f(x)$ be an integrable* (and hence bounded) function defined in the interval $-\pi \leq x \leq \pi$, and let $S_n(x)$ be the trigonometric polynomial

$$\frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

where the coefficients a_k and b_k are to be selected so as to render the integral

$$(101\text{-}1)\quad I_n = \int_{-\pi}^{\pi} \left[f(x) - \frac{a_0}{2} - \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \right]^2 dx$$

a minimum.

---

* See p. 110.

Calculating the partial derivatives gives

$$(101\text{-}2)\quad\begin{cases}\dfrac{\partial I_n}{\partial a_0} = -\displaystyle\int_{-\pi}^{\pi}\left[f(x) - \frac{a_0}{2} - \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)\right]dx,\\[2ex] \dfrac{\partial I_n}{\partial a_k} = -2\displaystyle\int_{-\pi}^{\pi}\left[f(x) - \frac{a_0}{2} - \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)\right]\cos kx\,dx,\\ \hfill (k = 1, 2, \cdots, n),\\[2ex] \dfrac{\partial I_n}{\partial b_k} = -2\displaystyle\int_{-\pi}^{\pi}\left[f(x) - \frac{a_0}{2} - \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)\right]\sin kx\,dx,\\ \hfill (k = 1, 2, \cdots, n).\end{cases}$$

But it is known that

$$(101\text{-}3)\quad\begin{cases}\displaystyle\int_{-\pi}^{\pi}\cos kx\cos mx\,dx = 0, & \text{if}\quad k \neq m;\\ \qquad\qquad = \pi, & \text{if}\quad k = m \neq 0;\\ \displaystyle\int_{-\pi}^{\pi}\sin kx\cos mx\,dx = 0; & \\ \displaystyle\int_{-\pi}^{\pi}\sin kx\sin mx\,dx = 0, & \text{if}\quad k \neq m;\\ \qquad\qquad = \pi, & \text{if}\quad k = m \neq 0;\end{cases}$$

so that carrying out the integrations indicated in (101-2) gives

$$(101\text{-}4)\quad\begin{cases}\dfrac{\partial I_n}{\partial a_0} = -\displaystyle\int_{-\pi}^{\pi}f(x)\,dx + \pi a_0;\\[2ex] \dfrac{\partial I_n}{\partial a_k} = -2\displaystyle\int_{-\pi}^{\pi}f(x)\cos kx\,dx + 2\pi a_k,\\ \hfill (k = 1, 2, \cdots, n);\\[2ex] \dfrac{\partial I_n}{\partial b_k} = -2\displaystyle\int_{-\pi}^{\pi}f(x)\sin kx\,dx + 2\pi b_k,\\ \hfill (k = 1, 2, \cdots, n).\end{cases}$$

The necessary condition for a minimum of I_n (see Sec. 89) requires that these derivatives vanish, and this leads to the following values for the coefficients:

$$(101\text{-}5)\quad\begin{cases}a_k = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(x)\cos kx\,dx, & (k = 0, 1, 2, \cdots, n);\\[2ex] b_k = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(x)\sin kx\,dx, & (k = 1, 2, \cdots, n).\end{cases}$$

It follows immediately from (101-4) that the coefficients a_k and b_k so obtained make (101-1) a minimum, since

$$\frac{\partial^2 I_n}{\partial a_k^2} = \frac{\partial^2 I_n}{\partial b_k^2} = 2\pi, \qquad (k = 1, 2, \cdots, n),$$
$$\frac{\partial^2 I_n}{\partial a_0^2} = \pi,$$

whereas all the mixed derivatives of the second order vanish. The constants a_k and b_k defined by the formulas (101-5) are called the *Fourier coefficients of the function* $f(x)$.

Some insight into the character of the approximation of a function $f(x)$ by the trigonometric polynomial can be gained by expanding the integrand in (101-1) and integrating those terms which are free of $f(x)$. The result is

$$\begin{aligned} I_n = {} & \int_{-\pi}^{\pi} [f(x)]^2\,dx - a_0 \int_{-\pi}^{\pi} f(x)\,dx \\ & - 2\sum_{k=1}^{n} \left[a_k \int_{-\pi}^{\pi} f(x)\cos kx\,dx + b_k \int_{-\pi}^{\pi} f(x)\sin kx\,dx \right] \\ & + \frac{\pi a_0^2}{2} + \pi \sum_{k=1}^{n} (a_k^2 + b_k^2), \end{aligned}$$

which, upon substitution from (101-5), gives the important formula for the magnitude of the integral of the square of the deviation,

$$\text{(101-6)} \qquad I_n = \int_{-\pi}^{\pi} [f(x)]^2\,dx - \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{n} (a_k^2 + b_k^2)\right].$$

Since I_n is never negative and the sum $\sum_{k=1}^{n} (a_k^2 + b_k^2)$ is an increasing function of n, it follows that I_n is a monotone-decreasing function of n. Consequently, the approximation .of $f(x)$ by the trigonometric polynomial

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

where the coefficients are determined by (101-5), improves with the increase in the number of terms of the polynomial.

Moreover, it follows directly from (101-6) that the series of the squares of the Fourier coefficients,

$$\text{(101-7)} \qquad \frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2),$$

converges whenever $f(x)$ is an integrable function. For, since $I_n \geq 0$, it is evident that

$$\text{(101-8)} \qquad \frac{a_0^2}{2} + \sum_{k=1}^{n} (a_k^2 + b_k^2) \leq \frac{1}{\pi}\int_{-\pi}^{\pi} [f(x)]^2 \, dx.$$

Now the left-hand member of the inequality (101-8) is precisely the nth partial sum of the series (101-7), and since it is bounded, the convergence of the series (101-7) follows directly from Theorem 1, Sec. 62. Furthermore, the terms of any convergent series form a null sequence, so that the Fourier coefficients of any integrable function $f(x)$ tend to zero with increasing n.

It was shown* by Liapunoff, in 1896, that the limit of I_n as n increases indefinitely is zero, so that the following remarkable formula is true whenever $f(x)$ is an integrable function.

$$\frac{1}{\pi}\int_{-\pi}^{\pi} [f(x)]^2 \, dx = \frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2).$$

However, the fact that the integral of the square of the deviation tends to zero when the number of terms in the trigonometric polynomial is increased indefinitely does not establish the convergence of the Fourier series

$$\text{(101-9)} \qquad \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

In fact, if $f(x)$ is merely an integrable function, the Fourier series may fail to converge either at some points of the interval or even in the entire interval $(-\pi, \pi)$. Even if the Fourier series (101-9) does converge, its sum is not necessarily $f(x)$.

---

* See WHITTAKER, E. T., and WATSON, G. N., Modern Analysis, p. 170.

In order to establish the convergence of (101-9), it is necessary to impose some additional restrictions on the function $f(x)$. There are several sets of sufficient conditions which will ensure convergence of the Fourier series. One of the most celebrated of these is due to Dirichlet, and it is set forth in the following section. The discussion of the next section makes use of the following theorem.

Theorem. *If $f(x)$ is an integrable function in the interval $\alpha \leq x \leq \beta$, then the integrals* $\int_\alpha^\beta f(x) \cos nx\, dx$ and $\int_\alpha^\beta f(x) \sin nx\, dx$ *tend to zero as n increases.*

If the interval (α, β) is of length 2π, then it follows from (101-5) that*

$$\int_\alpha^\beta f(x) \cos nx\, dx = \int_{-\pi}^{\pi} f(x) \cos nx\, dx = a_n \pi,$$

and it was noted above that the Fourier coefficients of an integrable function form a null sequence. If the interval (α, β) is of length less than 2π, define the function $\varphi(x)$ in the following way. Let $\varphi(x) \equiv f(x)$ in the interval (α, β), and let $\varphi(x) \equiv 0$ at all other points of the interval of length 2π. Then

$$\int_\alpha^\beta f(x) \cos nx\, dx = \int_{-\pi}^{\pi} \varphi(x) \cos nx\, dx = a_n \pi,$$

and it follows as before that $a_n \to 0$ as $n \to \infty$. If (α, β) is an interval of length greater than 2π, then it can be broken up into a finite number of intervals of lengths less than or equal to 2π. Then the integral

$$\int_\alpha^\beta f(x) \cos nx\, dx$$

can be expressed as the sum of a finite number of integrals of the type considered above, each of which tends to zero as $n \to \infty$.

The proof for the integral

$$\int_\alpha^\beta f(x) \sin nx\, dx$$

is entirely similar.

---

* It will be recalled that the function $f(x)$ is assumed to be periodic of period 2π.

102. Conditions of Dirichlet.

Theorem. *Let $f(x)$ be a function defined arbitrarily in the interval $-\pi \leqq x \leq \pi$, and outside this interval defined by the equation $f(x + 2\pi) = f(x)$. If $f(x)$ has a finite number of points of ordinary discontinuity and a finite number of maxima and minima in the interval $-\pi \leq x \leq \pi$, then it can be represented by the series*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

with

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \cos kt \cdot dt,$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \sin kt\, dt, \qquad (k = 0, 1, 2, \cdots),$$

*which converges at every point $x = x_0$ of the interval to the value**

$$\frac{f(x_0+) + f(x_0-)}{2}.$$

The restrictions imposed upon the function $f(x)$ in this theorem are known as the *Dirichlet conditions.*

In order to establish the theorem it is necessary to show that for any x in the interval

$$\lim_{n\to\infty} S_n(x) = \frac{f(x+) + f(x-)}{2},$$

where

$$(102\text{-}1) \qquad S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx).$$

---

* If $f(x)$ is continuous at the point $x = x_0$, then $f(x_0+) = f(x_0-) = f(x_0)$, so that at all points of continuity the series converges to $f(x)$. At the points of ordinary discontinuity it converges to the arithmetic mean of the values of the right- and left-hand limits.

Substituting the values of a_k and b_k in (102-1) gives

$$\begin{aligned} S_n(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{k=1}^{n} (\cos kt \cos kx + \sin kt \sin kx)\right] dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{k=1}^{n} \cos k(t - x)\right] dt. \end{aligned}$$

The expression appearing in the bracket can be summed with the aid of the trigonometric identity*

$$\frac{1}{2} + \sum_{k=1}^{n} \cos ku = \frac{\sin \dfrac{(2n + 1)u}{2}}{2 \sin \dfrac{u}{2}},$$

so that $S_n(x)$ can be written as

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \frac{\sin (2n + 1)\dfrac{t - x}{2}}{2 \sin \dfrac{t - x}{2}}\, dt.$$

Setting $t - x = 2u$, gives

$$\text{(102-2)} \qquad S_n(x) = \frac{1}{\pi}\int_{\frac{-\pi-x}{2}}^{\frac{\pi-x}{2}} f(x + 2u)\frac{\sin (2n + 1)u}{\sin u}\, du.$$

Thus the study of the limit of $S_n(x)$ as $n \to \infty$ is reduced to the study of the behavior of an integral of the form

---

[*] *Proof:*

$$\begin{aligned} \sin \frac{u}{2}\left(\frac{1}{2} + \cos u + \cos 2u + \cdots + \cos nu\right) \\ &= \frac{1}{2} \sin \frac{u}{2} + \sum_{k=1}^{n} \frac{1}{2}\left[\sin \left(k + \frac{1}{2}\right)u - \sin \left(k - \frac{1}{2}\right)u\right] \\ &= \frac{\sin \left(n + \frac{1}{2}u\right)}{2}. \end{aligned}$$

$$(102\text{-}3) \qquad \int_a^b f(u)\frac{\sin mu}{\sin u}\,du,$$

when m is allowed to increase indefinitely. This integral is known as the *integral of Dirichlet*, and a study of it reveals that the limiting value of (102-2) as $n \to \infty$ is precisely equal to $\frac{1}{2}[f(x+) + f(x-)]$. Unfortunately, a detailed discussion of the integral (102-3) is very involved,* and for this reason the convergence of (102-2) will be established under the hypothesis that $f(x)$ is a differentiable, and hence, continuous, function.

Consider first the function $F(x) = 1$. All the Fourier coefficients, with the exception of a_0, of such a function vanish, so that (102-2) gives the identity

$$1 = \frac{1}{\pi}\int_{\frac{-\pi-x}{2}}^{\frac{\pi-x}{2}} \frac{\sin(2n+1)u}{\sin u}\,du,$$

which is valid for all values of n. If both sides of this identity are multiplied by $f(x)$, there results

$$(102\text{-}4) \qquad f(x) = \frac{1}{\pi}\int_{\frac{-\pi-x}{2}}^{\frac{\pi-x}{2}} f(x)\frac{\sin(2n+1)u}{\sin u}\,du.$$

Subtracting (102-4) from (102-2) gives

$$(102\text{-}5) \qquad S_n(x) - f(x) = \frac{1}{\pi}\int_{\frac{-\pi-x}{2}}^{\frac{\pi-x}{2}} [f(x+2u) - f(x)]\frac{\sin(2n+1)u}{\sin u}\,du$$

$$= \frac{1}{\pi}\int_{\frac{-\pi-x}{2}}^{\frac{\pi-x}{2}} \left[\frac{f(x+2u) - f(x)}{2u}\cdot\frac{2u}{\sin u}\right]\sin(2n+1)u\,du.$$

Now the expression appearing in the bracket of the last integral is a continuous function of u, except, possibly, when $u = 0$. But

$$\lim_{u\to 0}\frac{f(x+2u) - f(x)}{2u}\cdot\frac{2u}{\sin u} = 2f'(x),$$

---

* Knopp, K., Theory and Application of Infinite Series, p. 356.
Carslaw, H. S., Fourier's Series and Integrals, p. 207.

since

$$\lim_{u\to 0}\frac{u}{\sin u} = 1.$$

Hence, if the value of the function

$$F(u) \equiv \frac{f(x+2u)-f(x)}{2u}\cdot\frac{2u}{\sin u}$$

is defined at $u = 0$ by the equation

$$F(0) \equiv 2f'(x),$$

then $F(u)$ is surely bounded. But the theorem of the preceding section states that for any integrable function $\varphi(u)$

$$\lim_{n\to\infty}\int_a^b \varphi(u)\sin nu\,du = 0,$$

so that the integral in (102-5) approaches zero as $n \to \infty$. Accordingly,

$$\lim_{n\to\infty} S_n(x) = f(x).$$

The foregoing proof can be extended to include the case of a function with a finite number of ordinary discontinuities if the function possesses a derivative at every point where it is continuous.

103. Orthogonal Functions. A set of continuous functions

$$(103\text{-}1)\qquad u_1(x),\ u_2(x),\ \cdots,\ u_n(x),\ \cdots,$$

which do not vanish identically in the interval $a \leq x \leq b$, is said to be orthogonal with respect to the interval (a, b) if the functions $u_i(x)$ satisfy the relations

$$(103\text{-}2)\qquad \int_a^b u_i(x)u_j(x)\,dx = 0,\qquad \text{if}\qquad i \neq j.$$

For $i = j$, (103-2) becomes

$$\int_a^b [u_i(x)]^2\,dx \equiv c_i^2,$$

where c_i^2 certainly is not zero.

If each of the orthogonal functions $u_i(x)$ be divided by c_i there will be obtained a system of *normal* orthogonal functions,

$$v_1(x) = \frac{u_1(x)}{c_1}, \qquad v_2(x) = \frac{u_2(x)}{c_2}, \; \cdots, \; v_n(x) = \frac{u_n(x)}{c_n}, \; \cdots,$$

which are characterized by the property that

$$\begin{aligned} \int_a^b v_i(x)v_j(x)\,dx &= 0, \qquad \text{if} \qquad i \neq j, \\ &= 1, \qquad \text{if} \qquad i = j. \end{aligned} \tag{103-3}$$

Consider a set of normal orthogonal functions $v_i(x)$, and assume that an arbitrary function $f(x)$ can be expanded in a series

$$\begin{aligned} f(x) &= a_1v_1(x) + a_2v_2(x) + \cdots + a_nv_n(x) + \cdots \\ &\equiv \sum_{i=1}^{\infty} a_iv_i(x), \end{aligned} \tag{103-4}$$

which can be integrated term by term. Multiplying both sides of (103-4) by $v_j(x)$, and integrating term by term between the limits a and b,

$$\int_a^b f(x)v_j(x)\,dx = \sum_{i=1}^{\infty} a_i \int_a^b v_i(x)v_j(x)\,dx,$$

which, by virtue of (103-3), gives the formula

$$a_i = \int_a^b f(x)v_i(x)\,dx, \qquad (i = 1, 2, 3, \cdots). \tag{103-5}$$

The numbers a_i are known as the Fourier coefficients of the function $f(x)$ associated with the system of normal and orthogonal functions

$$v_1(x), v_2(x), \cdots, v_n(x), \cdots.$$

The set of functions

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \cdots$$

is obviously a normal orthogonal set in the interval $(-\pi, \pi)$.

Instead of approaching the subject of the expansion of arbitrary functions in a trigonometric series with the aid of the criterion of Sec. 100, one could begin by assuming that it is possible to expand the function $f(x)$ in a series

$$(103\text{-}6)\qquad f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx),$$

which can be integrated term by term in the interval $(-\pi, \pi)$. Multiplying both sides of (103-6) by $\cos mx$, and integrating term by term with the aid of the formulas (101-3), gives

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos kx\, dx, \qquad (k = 0, 1, 2, \cdots).$$

Similarly, multiplying (103-6) by $\sin mx$, and integrating the resulting series term-wise, furnishes

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin kx\, dx, \qquad (k = 1, 2, 3, \cdots).$$

Thus, the hypothesis that the function $f(x)$ can be expanded in a trigonometric series (103-6) that can be integrated term by term leads to the same values of the coefficients as those obtained earlier by an entirely different method.

It is seen from the foregoing that the discussion of Secs. 100 and 101 constitutes a very special case of the problem of representing an arbitrary function $f(x)$ by a series of orthogonal functions.

104. Expansion of Functions in Fourier Series. This section contains some illustrative examples of expansion of functions, satisfying the Dirichlet conditions in the interval $(-\pi, \pi)$, in the series

$$(104\text{-}1)\qquad \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

where the coefficients a_n and b_n are given by the formulas

$$(104\text{-}2)\qquad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx\, dx,$$

and

$$(104\text{-}3)\qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx\, dx.$$

Illustrative Example 1. Expand $f(x) = x$ in Fourier series in the interval $-\pi \leqq x \leqq \pi$. Calculating the coefficients a_n and b_n gives

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \cos nx\, dx = 0,$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx\, dx = -\frac{2}{n}\cos n\pi.$$

Hence,

$$x = 2[(-\tfrac{1}{1}\cos \pi)\sin x + (-\tfrac{1}{2}\cos 2\pi)\sin 2x + (-\tfrac{1}{3}\cos 3\pi)\sin 3x + \cdots]$$

or

$$x = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right).$$

In this particular case only the sine terms remain. It may be noted that whenever the function $f(x)$ is an odd function, that is, when $f(-x) = -f(x)$, then $a_n = 0$, for $n = 0, 1, 2, \cdots$, since, for such a function,

$$\int_{-\pi}^{0} f(x)\cos nx\, dx = -\int_{0}^{\pi} f(x)\cos nx\, dx.$$

Similarly, if $f(x)$ is an even function, that is, when $f(-x) = f(x)$, then $b_n = 0$, for $n = 1, 2, 3, \cdots$, since

$$\int_{-\pi}^{0} f(x)\sin nx\, dx = -\int_{0}^{\pi} f(x)\sin nx\, dx,$$

so that the function would be represented by a series of cosine terms.

If in the foregoing illustration the first four terms be plotted by composition of

$$y = 2\sin x,\quad y = -\sin 2x,\quad y = \tfrac{2}{3}\sin 3x,\quad y = -\tfrac{1}{2}\sin 4x,$$

the curve

$$y = 2\sin x - \sin 2x + \tfrac{2}{3}\sin 3x - \tfrac{1}{2}\sin 4x$$

is obtained. It is represented on Fig. 83. As the number of terms is increased, the approximating curves approach $y = x$ as a limit for all values of x, $-\pi < x < \pi$, but not for $x = \pm\pi$. Since the series has period 2π, it represents the discontinuous function, shown in Fig. 84 by a series of parallel lines. It should be noted that each term of the series is continuous and the function from which the series was derived is continuous, but the function represented by the series has finite discontinuities at

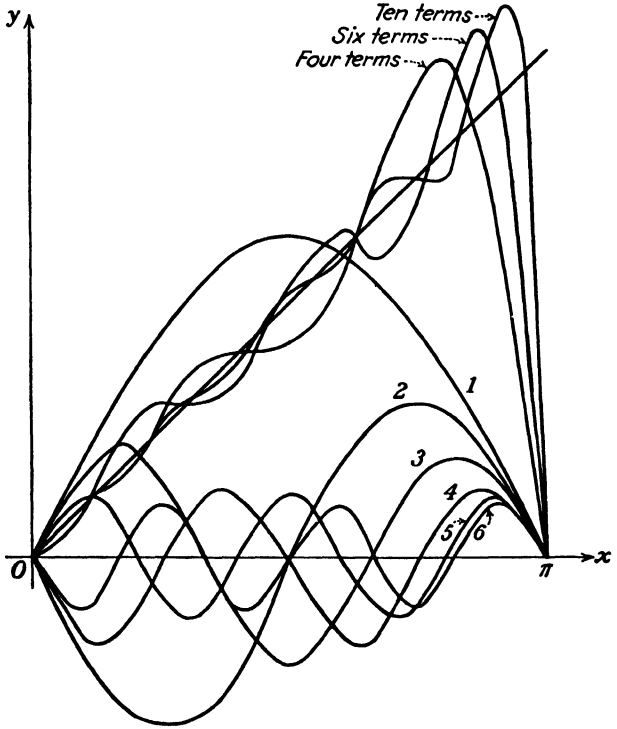

Fig. 83.

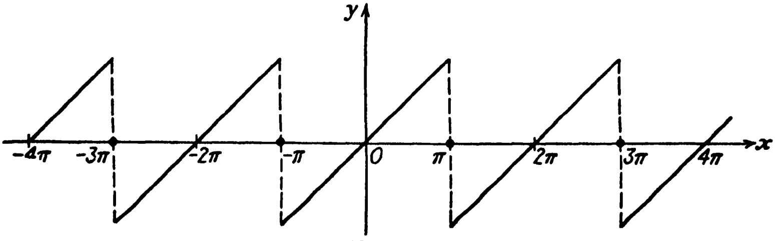

Fig 84.

$x = \pm(2k+1)\pi$. At such points the series converges to zero which is one-half the value of the sum of the right- and left-hand limits.

Illustrative Example 2. Develop $f(x)$ in Fourier series in the interval $(-\pi, \pi)$, if

$$\begin{aligned} f(x) &= 0, &&\text{for} && -\pi < x < 0, \\ &= \pi, &&\text{for} && 0 < x < \pi. \end{aligned}$$

Now

$$\begin{aligned} a_0 &= \frac{1}{\pi}\left[\int_{-\pi}^{0} 0 \cdot dx + \int_{0}^{\pi} \pi\, dx\right] = \pi, \\ a_n &= \frac{1}{\pi}\int_0^{\pi} \pi \cos nx\, dx = 0, \\ b_n &= \frac{1}{\pi}\int_0^{\pi} \pi \sin nx\, dx = \frac{1}{n}(1 - \cos n\pi). \end{aligned}$$

The series is then

$$\frac{\pi}{2} + 2\left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right)$$

The graph of $f(x)$ from $-\pi$ to π consists of the x-axis from $-\pi$ to 0, and the line AB from 0 to π (see Fig. 85). There is a finite discontinuity for $x = 0$. For $x = 0$ the series reduces to $\frac{\pi}{2}$, which is equal to half the sum of $\lim\limits_{\epsilon \to 0} f(0 - \epsilon)$ and $\lim\limits_{\epsilon \to 0} f(0 + \epsilon)$.

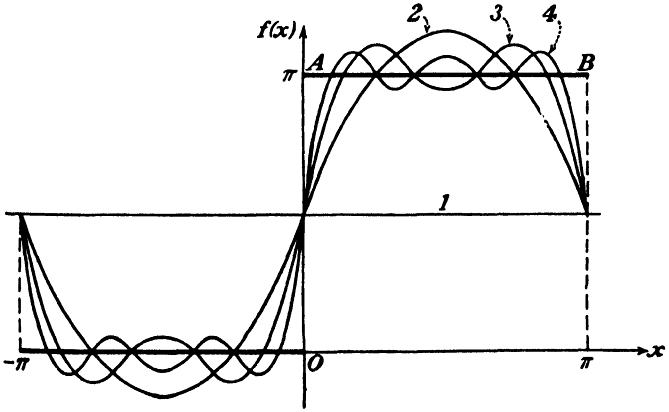

FIG. 85.

It may be observed from the series that every approximation curve will pass through the point $\left(0, \frac{\pi}{2}\right)$. The figure shows the first, second, third, and fourth approximation curves whose equations are

$$y = \frac{\pi}{2}, \qquad y = \frac{\pi}{2} + 2 \sin x, \qquad y = \frac{\pi}{2} + 2\left(\sin x + \frac{\sin 3x}{3}\right),$$
$$y = \frac{\pi}{2} + 2\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5}\right),$$

as well as the graph of $f(x)$.

At $x = \pm\pi$ the series reduces to $\frac{\pi}{2}$, and again every approximation curve gives this same value for the ordinate at $\pm\pi$. This value is one half the sum of $f(-\pi+)$ and $f(\pi-)$.

Illustrative Example 3. Let $f(x)$ be defined by the relations

$$\begin{aligned} f(x) &= -\pi, \quad &&\text{if} \quad -\pi < x < 0,\\ &= x, &&\text{if} \quad 0 < x < \pi; \end{aligned}$$

then the Fourier coefficients for $f(x)$ are given by

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} -\pi\,dx + \int_0^{\pi} x\,dx\right] = \frac{1}{\pi}\left(-\pi^2 + \frac{\pi^2}{2}\right)\\ &= -\frac{\pi}{2},\\ a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} -\pi\cos nx\,dx + \int_0^{\pi} x\cos nx\,dx\right]\\ &= \frac{1}{\pi}\left[0 + \frac{\cos n\pi}{n^2} - \frac{1}{n^2}\right]\\ &= \frac{1}{\pi}\left[\frac{\cos n\pi - 1}{n^2}\right],\\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} -\pi\sin nx\,dx + \int_0^{\pi} x\sin nx\,dx\right]\\ &= \frac{1}{\pi}\left[\frac{\pi}{n} - \frac{\pi}{n}\cos n\pi - \frac{\pi}{n}\cos n\pi\right]\\ &= \frac{1}{n}(1 - 2\cos n\pi). \end{aligned}$$

Therefore,

$$f(x) = -\frac{\pi}{4} - \frac{2}{\pi}\cos x - \frac{2}{\pi}\frac{\cos 3x}{3^2} - \frac{2}{\pi}\frac{\cos 5x}{5^2} - \cdots$$
$$+ 3\sin x - \frac{\sin 2x}{2} + \frac{3\sin 3x}{3} - \frac{\sin 4x}{4} + \frac{3\sin 5x}{5} - \cdots.$$

When $x = 0$, the series reduces to

$$-\frac{\pi}{4} - \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right),$$

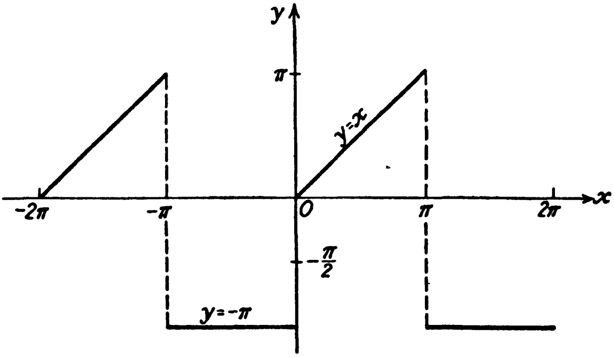

Fig. 86.

which must coincide with (see Fig. 86)

$$\frac{f(0+) + f(0-)}{2} = -\frac{\pi}{2}.$$

Thus,

$$-\frac{\pi}{4} - \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) = -\frac{\pi}{2}.$$

Hence,

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}.$$

Also for $x = \pm\pi$, the series gives

$$-\frac{\pi}{4} + \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right) = 0,$$

since

$$\frac{f(-\pi+)+f(\pi-)}{2}=0.$$

This example suggests the use of Fourier series in evaluating sums of series of constants.

PROBLEMS

1. Show that

$$x^2=\frac{\pi^2}{3}+4\sum_{n=1}^{\infty}(-1)^n\frac{\cos nx}{n^2},\qquad (-\pi \leqq x \leq \pi).$$

2. If

$$\begin{aligned} f(x) &= -x, &&\text{for}\quad -\pi<x<0,\\ &= 0, &&\text{for}\quad 0<x<\pi; \end{aligned}$$

then

$$f(x)=\frac{\pi}{4}-\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2}+\sum_{n=1}^{\infty}\frac{(-1)^n\sin nx}{n}.$$

3. If

$$\begin{aligned} f(x) &= 0, &&\text{for}\quad -\pi\leq x\leq 0,\\ &= \sin x, &&\text{for}\quad 0\leq x\leq\pi; \end{aligned}$$

then

$$f(x)=\frac{1}{\pi}-\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\cos 2nx}{4n^2-1}-\frac{1}{2}\sin x.$$

4. If $f(x)=e^x$ in the interval $(0, 2\pi)$, then

$$e^x=\frac{e^{2\pi}-1}{\pi}\left[\frac{1}{2}+\sum_{n=1}^{\infty}\frac{\cos nx}{1+n^2}-\sum_{n=1}^{\infty}\frac{n\sin nx}{1+n^2}\right].$$

5. Deduce from Prob. 1 that

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^2}=\frac{\pi^2}{12}.$$

6. Show that

$$\cos \alpha x = \frac{\sin \pi\alpha}{\pi\alpha} + \sum_{n=1}^{\infty} (-1)^n \frac{2\alpha \sin \pi\alpha}{\pi(\alpha^2 - n^2)} \cos nx,$$

if $-\pi \leqq x \leqq \pi$.

7. Deduce from Prob. 6 that

$$\cot \pi\alpha = \frac{1}{\pi}\left(\frac{1}{\alpha} - \sum_{n=1}^{\infty} \frac{2\alpha}{n^2 - \alpha^2}\right).$$

8. Deduce from the expansion of $f(x) = x + x^2$ in Fourier series in the interval $(-\pi, \pi)$ that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

9. Expand $x \sin x$ and $x \cos x$ in Fourier series in the interval $(0, 2\pi)$.

10. Find the Fourier series expansion for $f(x)$, if

$$\begin{aligned} f(x) &= \frac{\pi}{2}, \qquad \text{for} \qquad -\pi < x < \frac{\pi}{2}, \\ &= 0, \qquad \text{for} \qquad \frac{\pi}{2} < x < \pi. \end{aligned}$$

105. Sine and Cosine Series. The Fourier expansion for $f(x)$ in $(-\pi, \pi)$ has the form (104-1), in which the coefficients a_n and b_n are given by (104-2) and (104-3). As previously observed (Sec. 104), if $f(x)$ is an even function, (104-1) reduces to a series containing only cosine terms; and if $f(x)$ is an odd function, (104-1) reduces to a series containing only sine terms. Now suppose that it is desired that $f(x)$ be expanded in a Fourier series which will be used for the interval 0 to π only. In that case it is frequently convenient to obtain the expansion in terms of sines alone or in terms of cosines alone. For this purpose define

$$F(x) \equiv f(x) \qquad \text{for} \qquad 0 < x < \pi,$$

and

$$F(x) \equiv f(-x) \qquad \text{for} \qquad -\pi < x < 0,$$

so that $F(x)$ is an even function identical with $f(x)$ in $0 < x < \pi$.

For an even function:

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\sin nx\,dx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} F(x)\sin nx\,dx + \int_{0}^{\pi} F(x)\sin nx\,dx\right] \\ &= \frac{1}{\pi}\left[-\int_{0}^{\pi} F(-x)\sin(-nx)(-dx) + \int_{0}^{\pi} F(x)\sin nx\,dx\right] \\ &= \frac{1}{\pi}\left[-\int_{0}^{\pi} F(x)\sin nx\,dx + \int_{0}^{\pi} F(x)\sin nx\,dx\right] = 0, \end{aligned}$$

and

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\cos nx\,dx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} F(x)\cos nx\,dx + \int_{0}^{\pi} F(x)\cos nx\,dx\right] \\ &= \frac{1}{\pi}\left[\int_{0}^{\pi} F(x)\cos nx\,dx + \int_{0}^{\pi} F(x)\cos nx\,dx\right] \\ &= \frac{2}{\pi}\int_{0}^{\pi} F(x)\cos nx\,dx. \end{aligned}$$

Hence, in the expansion of $F(x)$ in the interval $-\pi < x < \pi$ only the cosine terms appear. Moreover, $F(x)$ is identical with $f(x)$ for $0 < x < \pi$. Therefore,*

$$(105\text{-}1)\qquad f(x) = \frac{a_0}{2} + a_1\cos x + a_2\cos 2x + \cdots + a_n\cos nx + \cdots$$

in the interval $(0, \pi)$, where

$$(105\text{-}2)\qquad a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx\,dx.$$

Similarly, if $F(x)$ be defined so that

$$F(x) \equiv f(x) \qquad \text{for} \qquad 0 < x < \pi,$$

and

$$F(x) \equiv -f(-x) \qquad \text{for} \qquad -\pi < x < 0,$$

---

* If $f(x)$ has a finite discontinuity at the point $x = x_0$, then the left-hand member of (105-1) is defined to be $\frac{1}{2}[f(x_0+) + f(x_0-)]$.

then the a_n all vanish and

$$f(x) = b_1 \sin x + b_2 \sin 2x + \cdots + b_n \sin nx + \cdots, \tag{105-3}$$

where

$$b_n = \frac{2}{\pi}\int_0^\pi f(x) \sin nx\, dx. \tag{105-4}$$

Thus $f(x)$ can be represented in the interval $0 < x < \pi$ by either (105-1) or (105-3). Frequently one series is more desirable than the other.

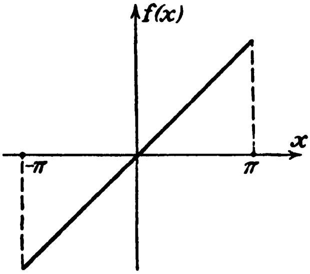

Fig. 87.

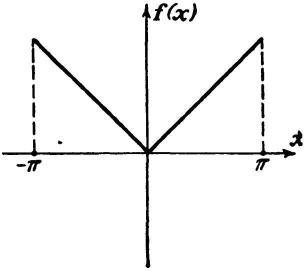

Fig. 88.

Example. As has been determined already (see Example 1, Sec. 104), the expansion for $f(x) = x$ in a sine series is

$$x = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right).$$

This series represents $f(x) = x$ in the interval $(-\pi, \pi)$. If one is interested in the values of the function in the interval $(0, \pi)$, the same function can be expanded in a series of cosines.

In fact, in the interval $(0, \pi)$,

$$x = \frac{\pi}{2} - \frac{4}{\pi}\left(\frac{\cos x}{1} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right),$$

since

$$a_0 = \frac{2}{\pi}\int_0^\pi x\, dx = \pi,$$

and

$$a_n = \frac{2}{\pi}\int_0^\pi x \cos nx\, dx = \frac{2}{n^2\pi}[(-1)^n - 1].$$

The sine series represents the odd function shown in Fig. 87 and the cosine series the even function in Fig. 88. The two graphs are identical in the interval $(0, \pi)$.

PROBLEMS

1. Show that if c is a constant, then, in $0 < x < \pi$,

$$c = c\frac{4}{\pi}\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right).$$

2. Give sine and cosine developments of $y = x \sin x$ in the interval $(0, \pi)$.

3. Show that, in $(0, \pi)$,

$$x^2 = \frac{2}{\pi}\left[\left(\frac{\pi^2}{1} - \frac{4}{1^3}\right)\sin x - \frac{\pi^2}{2}\sin 2x + \left(\frac{\pi^2}{3} - \frac{4}{3^3}\right)\sin 3x - \frac{\pi^2}{4}\sin 4x + \cdots\right].$$

4. Prove that if $f(x)$ is any function of x, it can be expressed as the sum of an even function of x and an odd function of x.

5. Show that if $f(x) = x$ for $0 < x < \frac{\pi}{2}$, and $f(x) = \pi - x$ for $\frac{\pi}{2} < x < \pi$, then

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\left(\frac{\cos 2x}{1^2} + \frac{\cos 6x}{3^2} + \frac{\cos 10x}{5^2} + \cdots\right).$$

6. Show that

$$\log\left(2 \sin \frac{x}{2}\right) = -\sum_{n=1}^{\infty} \frac{\cos nx}{n}, \quad \text{if} \quad 0 < x < \pi.$$

7. Find the expansion in the series of sines, if

$$\begin{aligned} f(x) &= \frac{\pi}{4}x, & 0 \le x \le \frac{\pi}{2}, \\ &= \frac{\pi}{4}(\pi - x), & \frac{\pi}{2} \le x \le \pi. \end{aligned}$$

8. Expand $f(x) = e^x$ in the series of cosines in the interval $(0, \pi)$.

106. Extension of Interval of Expansion. The methods developed up to this point restrict the interval in which $f(x)$ can be expanded in a Fourier series to $(-\pi, \pi)$. In many problems it is desirable to develop $f(x)$ in a Fourier series which will be valid over a wider interval. In order to obtain an expansion which will hold for the interval $(-l, l)$, change the variable by replacing x by $\frac{l}{\pi}z$. Then $f(x) = f\left(\frac{l}{\pi}z\right)$ can be developed in a Fourier series in z,

$$(106\text{-}1) \qquad f\left(\frac{l}{\pi}z\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nz + \sum_{n=1}^{\infty} b_n \sin nz,$$

in which

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \cos nz\, dz$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \sin nz\, dz.$$

The expression (106-1) will be valid for $-\pi < z < \pi$, but $z = \frac{\pi x}{l}$ so that (106-1) becomes

$$(106\text{-}2) \qquad f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{l} + \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{l}.$$

Also,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \cos nz\, dz = \frac{1}{l}\int_{-l}^{l} f(x) \cos \frac{n\pi x}{l}\, dx$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lz}{\pi}\right) \sin nz\, dz = \frac{1}{l}\int_{-l}^{l} f(x) \sin \frac{n\pi x}{l}\, dx.$$

Example. Develop $f(x)$ in Fourier series in the interval $(-2, 2)$, if $f(x) = 0$ for $-2 < x < 0$, and $f(x) = 1$ for $0 < x < 2$.

Here

$$a_0 = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot dx + \int_{0}^{2} 1 \cdot dx\right) = 1,$$

$$a_n = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot \cos\frac{n\pi x}{2}\,dx + \int_{0}^{2} 1 \cdot \cos\frac{n\pi x}{2}\,dx\right) = \frac{1}{n\pi}\sin\frac{n\pi x}{2}\Big|_0^2 = 0,$$

$$b_n = \frac{1}{2}\left(\int_{-2}^{0} 0 \cdot \sin\frac{n\pi x}{2}\,dx + \int_{0}^{2} 1 \cdot \sin\frac{n\pi x}{2}\,dx\right) = \frac{1}{n\pi}(1 - \cos n\pi).$$

Therefore,

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\left(\sin\frac{\pi x}{2} + \frac{1}{3}\sin\frac{3\pi x}{2} + \frac{1}{5}\sin\frac{5\pi x}{2} + \cdots\right).$$

PROBLEMS

1. The expansion of $f(x)$ is desired for $0 < x < l$. If $F(x) \equiv f(x)$ for $0 < x < l$ and $F(x) \equiv -f(-x)$ for $-l < x < 0$ (that is, $F(x)$ is defined as an odd function), show that the expansion of $F(x)$ and $f(x)$ for $0 < x < l$ is

$$\sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l},$$

where

$$b_n = \frac{2}{l}\int_0^l f(x)\sin\frac{n\pi x}{l}\,dx.$$

If $\varphi(x) \equiv f(x)$ for $0 < x < l$, and $\varphi(x) \equiv f(-x)$ for $-l < x < 0$ (that is, $\varphi(x)$ is defined as an even function), show that the expansion of $\varphi(x)$ and $f(x)$ for $0 < x < l$ is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l},$$

where

$$a_n = \frac{2}{l}\int_0^l f(x)\cos\frac{n\pi x}{l}\,dx.$$

2. Using the results of the preceding problem obtain the sine and cosine expansions of the following functions:

(*a*) $f(x) = 1$ in the interval $(0, 2)$;
(*b*) $f(x) = x$ in the interval $(0, 1)$;
(*c*) $f(x) = x^2$ in the interval $(0, 3)$.

3. Expand $f(x) = \cos \pi x$ in the interval $(-1, 1)$.

4. Expand

$$f(x) = \begin{cases} \tfrac{1}{4} - x, & \text{if} \quad 0 < x < \tfrac{1}{2}, \\ x - \tfrac{3}{4}, & \text{if} \quad \tfrac{1}{2} < x < 1, \end{cases}$$

in the series of sines.

5. Find the expansion in the series of cosines, if

$$f(x) = \begin{cases} 0, & \text{if} \quad 0 < x < 1, \\ 1, & \text{if} \quad 1 < x < 2. \end{cases}$$

6. Expand $f(x) = |x|$ in the series of cosines in the interval $(-1, 1)$.

7. Show that the series

$$\frac{l}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin \frac{2n\pi x}{l}$$

represents $\tfrac{1}{2}l - x$ when $0 < x < l$.

8. Find the expansion in the series of cosines, if

$$f(x) = \begin{cases} 1, & \text{when} \quad 0 < x < \pi, \\ 0, & \text{when} \quad \pi < x < 2\pi. \end{cases}$$

107. Complex Form of Fourier Series. The Fourier series

(107-1) $$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

with

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt,$$

can be written with the aid of the Euler formula*

(107-2) $$e^{iu} = \cos u + i \sin u$$

in an equivalent form, namely,

(107-3) $$f(x) = \sum_{n=-\infty}^{n=+\infty} c_n e^{inx},$$

where the coefficients c_n are defined by the equation

(107-4) $$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} \, dt.$$

---

* See Sec. 84.

The index of summation n in (107-3) runs through the set of all positive and negative integral values including zero.

The equivalence of (107-3) and (107-1) can be established in the following manner. Substituting from (107-2) in (107-4) gives for $n > 0$

$$\begin{aligned} c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\ (\cos nt - i \sin nt)\, dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \cos nt\, dt - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(t) \sin nt\, dt \\ &= \frac{a_n}{2} - i\frac{b_n}{2}. \end{aligned}$$

A similar calculation for $n < 0$ gives

$$c_{-n} = \frac{a_n}{2} + i\frac{b_n}{2},$$

while

$$c_0 = \frac{a_0}{2}.$$

Now (107-3) can be written in the form

$$f(x) = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=1}^{\infty} c_{-n} e^{-inx}.$$

Making use of the expressions for the c_n just found gives

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \frac{a_n - ib_n}{2} e^{inx} + \sum_{n=1}^{\infty} \frac{a_n + ib_n}{2} e^{-inx} \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \frac{e^{inx} + e^{-inx}}{2} - i\sum_{n=1}^{\infty} b_n \frac{e^{inx} - e^{-inx}}{2}. \end{aligned}$$

Recalling that

$$e^{iu} + e^{-iu} = 2 \cos u \qquad \text{and} \qquad e^{iu} - e^{-iu} = 2i \sin u,$$

gives

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

which establishes the identity of (107-3) with (107-1).

PROBLEM

Show that the Fourier series in the interval $(-l, l)$ can be written in the form

$$f(x) = \sum_{n=-\infty}^{n=+\infty} c_n e^{\frac{in\pi x}{l}},$$

where

$$c_n = \frac{1}{2l}\int_{-l}^{l} f(t)e^{-\frac{in\pi t}{l}}\,dt.$$

108. Differentiation and Integration of Fourier Series. It was proved in Sec. 101 that the Fourier coefficients of any integrable function defined in the interval $(-\pi, \pi)$ form a null sequence. A more precise bound for the magnitude of the Fourier coefficients can be obtained if the function $f(x)$ satisfies the conditions of Dirichlet in the interval $(-\pi, \pi)$. The requirement that $f(x)$ have only a finite number of maxima and minima is equivalent to saying that the interval $(-\pi, \pi)$ can be divided into a finite number of subintervals in each of which the function is monotone. Consequently,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx \tag{108-1}$$

can be expressed as the sum of a finite number of integrals of the form

$$\int_a^b f(x)\cos nx\,dx, \tag{108-2}$$

in each of which $f(x)$ is a monotone function.

The application of the second mean-value theorem for integrals (Sec. 37) to (108-2) gives*

$$\begin{aligned}\int_a^b f(x)\cos nx\,dx &= f(a+)\int_a^\xi \cos nx\,dx + f(b-)\int_\xi^b \cos nx\,dx\\ &= \frac{f(a+)(\sin n\xi - \sin na)}{n}\\ &\qquad + \frac{f(b-)(\sin nb - \sin n\xi)}{n}.\end{aligned}$$

---

* The symbols $f(a+)$ and $f(b-)$ are used instead of $f(a)$ and $f(b)$ since the function may be discontinuous at the end points.

Now the numerators of these fractions are bounded, so that

$$\left|\int_a^b f(x) \cos nx \, dx\right| < \frac{M}{n}.$$

Since (108-1) is expressible as the sum of a finite number of integrals of the form (108-2), it is clear that a_n is of order* $\frac{1}{n}$. An entirely similar argument can be carried through for b_n. This proves the following theorem:

Theorem 1. *The Fourier coefficients a_n and b_n of a function fulfilling the conditions of Dirichlet in the interval $(-\pi, \pi)$ satisfy the inequalities*

$$|a_n| < \frac{M}{n} \qquad \text{and} \qquad |b_n| < \frac{M}{n},$$

where M is a positive number independent of n.

If the function $f(x)$, defined in the interval $-\pi \leq x \leq \pi$, is periodic and has the derivative which satisfies the conditions of Dirichlet in the same interval, then the Fourier coefficients are of order $\frac{1}{n^2}$. For,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx \\ &= \frac{1}{n\pi}\Big[f(x) \sin nx\Big]_{-\pi}^{\pi} - \frac{1}{n\pi}\int_{-\pi}^{\pi} f'(x) \sin nx \, dx \\ &= -\frac{1}{n\pi}\int_{-\pi}^{\pi} f'(x) \sin nx \, dx. \end{aligned}$$

From Theorem 1 it follows that

$$\int_{-\pi}^{\pi} f'(x) \sin nx \, dx$$

is of order $\frac{1}{n}$, so that

$$|a_n| < \frac{M}{n^2}.$$

---

* $F(n)$ is said to be of order $\frac{1}{n^k}$ if $|F(n)| \leq \frac{M}{n^k}$, where M is independent of n. There are other definitions of the phrase *of the order of;* see, for example, Whittaker and Watson, Modern Analysis, p. 11.

A similar argument proves that

$$|b_n| < \frac{M}{n^2}.$$

This result can be extended readily to establish

Theorem 2. *If the function $f(x)$ is periodic and has $k - 1$ derivatives in the interval $-\pi \leq x \leq \pi$, and is such that its kth derivative satisfies the conditions of Dirichlet in the same interval, then the Fourier coefficients for $f(x)$ satisfy the inequalities*

$$|a_n| < \frac{M}{n^{k+1}} \qquad \text{and} \qquad |b_n| < \frac{M}{n^{k+1}},$$

where M is a positive number independent of n.

It should be noted that the requirement that the derivatives of the function $f(x)$ exist implies the continuity of the function. Moreover, since $f(x)$ is continuous in the interval $-\pi \leq x \leq \pi$, the periodicity condition requires that $f(-\pi) = f(\pi)$.

An important conclusion follows directly from Theorem 2. Consider a periodic function $f(x)$ which is continuous in the interval $(-\pi, \pi)$ and whose first derivative $f'(x)$ satisfies the conditions of Dirichlet in the same interval. The Fourier series for such a function has coefficients of order $\frac{1}{n^2}$, so that

$$\begin{aligned}|f(x)| &= \left|\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)\right| \\ &\leq \left|\frac{a_0}{2}\right| + \sum_{n=1}^{\infty}|a_n \cos nx + b_n \sin nx| \\ &< |a_0| + \sum_{n=1}^{\infty}\frac{M}{n^2},\end{aligned}$$

where M is a positive number independent of n.

This inequality states that the terms of the Fourier series are numerically less than the corresponding terms of a convergent series of positive constants. Hence, the series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

converges uniformly[*] in the interval $(-\pi, \pi)$. Accordingly, the Fourier series for a function whose derivative satisfies the Dirichlet conditions can be integrated term by term.

It is obvious that the differentiation of a trigonometric series

$$\sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

term by term, produces a trigonometric series whose coefficients are of the order of na_n and nb_n, so that the differentiated series will converge less rapidly if at all. For example, the series resulting from term-by-term differentiation of the series[†] representing the function $f(x) = x$, $(-\pi \leq x \leq \pi)$,

$$2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\right), \tag{108-3}$$

is

$$2(\cos x - \cos 2x + \cos 3x - \cdots),$$

which diverges for all values of x.

On the other hand, term-by-term integration of the same series between the limits 0 and x gives the series

$$\begin{aligned} S(x) &\equiv -2\left\{\frac{\cos x}{1^2}\Big|_0^x - \frac{\cos 2x}{2^2}\Big|_0^x + \frac{\cos 3x}{3^2}\Big|_0^x - \cdots\right\} \\ &= -2\left(\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - \cdots\right) \\ &\qquad\qquad + 2\left(1 - \frac{1}{2^2} + \frac{1}{3^2} - \cdots\right). \end{aligned}$$

But[‡]

$$\sum_{n=1}^{\infty} (-1)^{n-1}\frac{1}{n^2} = \frac{\pi^2}{12},$$

so that

$$S(x) = \frac{\pi^2}{6} - 2\sum_{n=1}^{\infty} (-1)^{n-1}\frac{\cos nx}{n^2},$$

---

[*] See Sec. 71.

[†] See Example 1, Sec. 104.

[‡] See Prob. 5, Sec. 104.

which is precisely the Fourier series for $\frac{x^2}{2}$ in the interval

$$-\pi \leq x \leq \pi.^*$$

Now despite the fact that the series (108-3) does not converge uniformly, the term-by-term integral of this series converges to the value of the integral of the function defined by (108-3). This is not an accident, and it is possible to prove that whenever $f(x)$ is represented by a Fourier series, so that

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

then

$$\int_a^x f(x)\, dx = \int_a^x \frac{a_0}{2}\, dx + \sum_{n=1}^{\infty} \int_a^x (a_n \cos nx + b_n \sin nx)\, dx.$$

In other words, the Fourier series representing $f(x)$ can always be integrated term by term regardless of whether the given series converges uniformly or not.†

However, great care must be exercised in differentiating the trigonometric series term by term. If $f(x)$ is periodic and continuous in the interval $-\pi \leq x \leq \pi$, and its derivative of the second order satisfies the Dirichlet conditions in the same interval, then the Fourier coefficients of $f(x)$ will be of order $\frac{1}{n^3}$. The series obtained by differentiating the Fourier series for $f(x)$ will have coefficients of order $\frac{1}{n^2}$, and consequently it will converge uniformly. Under such circumstances the term-by-term differentiation is clearly legitimate.

109. Fourier Integral. The theory developed in Sec. 106 permits one to write the Fourier series of a function $f(x)$ which is defined in the interval $(-l, l)$, where l is finite but can be made as large as desired. The limiting form of the series (106-2), when l is allowed to increase indefinitely, can be written in the form of an integral. It will be assumed here that the function $f(x)$ satisfies the conditions of Dirichlet in any interval $(-l, l)$, where l can be made as large as desired, and that the integral

---

* See Prob. 1, Sec. 104.

† The proof of this remarkable fact will be found in a paper by E. W Hobson in the *Journal of the London Mathematical Society*, vol. 2, p. 164, 1927 See also C. J. de la Vallée Poussin, Cours d'analyse infinitésimale. vol. 2.

$$\int_{-\infty}^{\infty} |f(x)|\, dx$$

converges.

Since $f(x)$ is assumed to satisfy the conditions of Dirichlet* in any interval $(-l, l)$, it follows that

$$\text{(109-1)} \qquad f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{l} + b_n \sin\frac{n\pi x}{l}\right),$$

where

$$a_n = \frac{1}{l}\int_{-l}^{l} f(t) \cos\frac{n\pi t}{l}\, dt, \qquad b_n = \frac{1}{l}\int_{-l}^{l} f(t) \sin\frac{n\pi t}{l}\, dt.$$

Substituting these values of the coefficients in (109-1) gives

$$\text{(109-2)} \quad f(x) = \frac{1}{2l}\int_{-l}^{l} f(t)\, dt + \frac{1}{l}\sum_{n=1}^{\infty}\int_{-l}^{l} f(t) \cos\frac{n\pi(t-x)}{l}\, dt.$$

Since $\int_{-\infty}^{\infty} |f(x)|\, dx$ is assumed to be convergent,

$$\left|\frac{1}{2l}\int_{-l}^{l} f(t)\, dt\right| \le \frac{1}{2l}\int_{-l}^{l} |f(t)|\, dt \le \frac{M}{2l},$$

which obviously tends to zero as l is allowed to increase indefinitely.

Now, if the interval $(-l, l)$ is made large enough, the quantity $\frac{\pi}{l}$, which appears in the integrands of the sum, can be made as small as desired. Therefore, the sum in (109-2) can be written as

---

* If $f(x)$ is not continuous at the point $x = x_0$, then the left-hand member of (109-1) means $\frac{1}{2}[f(x_0+) + f(x_0-)]$.

$$\begin{aligned}(109\text{-}3)\qquad &\frac{1}{\pi}\Big[\Delta\alpha\int_{-l}^{l} f(t)\cos\Delta\alpha(t-x)\,dt\\ &+\Delta\alpha\int_{-l}^{l} f(t)\cos 2\Delta\alpha(t-x)\,dt\\ &+\cdots\cdots\cdots\cdots\cdots\\ &+\Delta\alpha\int_{-l}^{l} f(t)\cos n\Delta\alpha(t-x)\,dt\\ &+\cdots\cdots\cdots\cdots\cdots\Big],\end{aligned}$$

where

$$\Delta\alpha = \frac{\pi}{l}.$$

This sum suggests the definition of the definite integral of the function

$$F(\alpha) = \int_{-l}^{l} f(t)\cos\alpha(t-x)\,dt,$$

in which the values of the function $F(\alpha)$ are calculated at the points

$$\frac{\pi}{l}, \frac{2\pi}{l}, \frac{3\pi}{l}, \ldots .$$

Now for large values of l

$$\int_{-l}^{l} f(t)\cos\alpha(t-x)\,dt$$

differs little from

$$\int_{-\infty}^{\infty} f(t)\cos\alpha(t-x)\,dt,$$

and it appears plausible that as l increases indefinitely the sum (109-3) will approach the limit

$$\frac{1}{\pi}\int_{0}^{\infty} d\alpha\int_{-\infty}^{\infty} f(t)\cos\alpha(t-x)\,dt.$$

If such is the case, then (109-2) can be written as

$$(109\text{-}4)\qquad f(x) = \frac{1}{\pi}\int_{0}^{\infty} d\alpha\int_{-\infty}^{\infty} f(t)\cos\alpha(t-x)\,dt.$$

The foregoing discussion is heuristic and cannot be regarded as a rigorous proof that $f(x)$ is capable of the integral representation (109-4). However, the validity of the formula (109-4) can be established rigorously,* if the function $f(x)$ satisfies the conditions enunciated above. The integral (109-4) bears the name of the *Fourier integral.*

The formula (109-4) assumes a simpler form if $f(x)$ is an even or an odd function. Thus, expanding the integrand of (109-4) gives

$$(109\text{-}5) \quad f(x) = \frac{1}{\pi}\int_0^\infty d\alpha \int_{-\infty}^\infty f(t) \cos \alpha t \cos \alpha x \, dt + \frac{1}{\pi}\int_0^\infty d\alpha \int_{-\infty}^\infty f(t) \sin \alpha t \sin \alpha x \, dt.$$

Now if $f(t)$ is an odd function, then the function $f(t) \cos \alpha t$ will be also odd. Therefore, for an odd function $f(x)$, (109-5) reduces to

$$f(x) = \frac{1}{\pi}\int_0^\infty d\alpha \int_{-\infty}^\infty f(t) \sin \alpha t \sin \alpha x \, dt,$$

since the first integral in (109-5) vanishes. Moreover, since $f(t) \sin \alpha t$ is an even function when $f(t)$ is odd, the foregoing integral can be written as

$$(109\text{-}6) \quad f(x) = \frac{2}{\pi}\int_0^\infty d\alpha \int_0^\infty f(t) \sin \alpha t \sin \alpha x \, dt.$$

An entirely similar argument proves that if $f(x)$ is an even function, then

$$(109\text{-}7) \quad f(x) = \frac{2}{\pi}\int_0^\infty d\alpha \int_0^\infty f(t) \cos \alpha t \cos \alpha x \, dt.$$

If $f(x)$ is defined only in the interval $(0, \pi)$, then either (109-6) or (109-7) can be used, since the function $f(x)$ may be thought to be defined in the interval $(-\pi, 0)$ so as to make it either even or odd. However, it must be remembered that at the points of discontinuity the integrals in (109-6) and (109-7) converge to $\frac{1}{2}[f(x_0+) + f(x_0-)]$, so that for $x = 0$ the formula (109-6) will always give the value 0.

---

* See Carslaw, H. S. Fourier's Series and Integrals, p. 283.

PROBLEMS

1. Set $f(x) = e^{-x}$ in (109-6) and show that

$$\int_0^\infty \frac{\alpha \sin \alpha x}{1 + \alpha^2}\, d\alpha = \frac{\pi}{2} e^{-x}, \qquad \text{if} \qquad x > 0.$$

2. Take

$$\begin{aligned} f(x) &= 1, &\quad \text{if} &\quad 0 \le x < 1; \\ &= 0, &\quad \text{if} &\quad x > 1 \end{aligned}$$

and show, with the aid of (109-7), that

$$\begin{aligned} \int_0^\infty \frac{\sin \alpha \cos \alpha x}{\alpha}\, d\alpha &= \frac{\pi}{2}, &\quad \text{if} &\quad 0 \le x < 1; \\ &= \frac{\pi}{4}, &\quad \text{if} &\quad x = 1; \\ &= 0, &\quad \text{if} &\quad x > 1. \end{aligned}$$

3. Show, with the aid of (109-6) and (109-7), that

$$\int_0^\infty \frac{\alpha \sin \alpha x}{\alpha^2 + \beta^2}\, d\alpha = \frac{\pi}{2} e^{-\beta x}, \qquad \text{if} \qquad \beta > 0;$$

$$\int_0^\infty \frac{\cos \alpha x}{\alpha^2 + \beta^2}\, d\alpha = \frac{\pi}{2\beta} e^{-\beta x}, \qquad \text{if} \qquad \beta > 0.$$

Hint: Take $f(x) = e^{-\beta x}$.

4. Show that the Fourier integral can be written in the form

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\alpha \int_{-\infty}^{\infty} f(t) e^{\alpha(t-x)i}\, dt.$$

5. Consider the equations

$$\int_0^\infty f(x) \cos xt\, dx = F(t)$$

and

$$\int_0^\infty \varphi(x) \sin xt\, dx = \Phi(t),$$

where $F(x)$ and $\Phi(x)$ are known functions and $f(x)$ and $\varphi(x)$ are the functions to be determined. The equations in which an unknown function appears under the integral sign are called *integral equations*. Show with the aid of (109-6) and (109-7) that

$$f(x) = \frac{2}{\pi} \int_0^\infty F(t) \cos xt\, dt$$

and

$$\varphi(x) = \frac{2}{\pi}\int_0^\infty \Phi(t) \sin xt \, dt$$

are the solutions of the given integral equations. Note particularly the remarkable form of these solutions. The integral equations of this problem are called the *integral equations of Fourier*.

CHAPTER XII

IMPLICIT FUNCTIONS

110. A Simple Problem in Implicit Functions. The purpose of this and the next two sections is to present a simplified formal treatment of some of the important problems arising in the theory of implicit functions. These problems are investigated in some detail in the subsequent sections where sharper definitions and restrictions on the type of functions to be considered are supplied.

A simple instance of an implicit functional relationship is that of the form

$$F(x, y) = 0, \tag{110-1}$$

where, corresponding to a certain range of real values of x, the equation (110-1) defines a set of real values of y. In particular, let (x_0, y_0) be a pair of values which satisfy (110-1), so that

$$F(x_0, y_0) \equiv 0,$$

and denote the totality of values of y, corresponding to a given range of values of x in the vicinity of $x = x_0$, by

$$y = f(x). \tag{110-2}$$

A question of practical importance will be considered next. Under what circumstances can one obtain a solution of (110-1) in terms of x in the neighborhood of (x_0, y_0)? If the function $f(x)$ can be expanded in Taylor's series about the point $x = x_0$, then

$$\begin{aligned} y = f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots \\ + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \cdots . \end{aligned} \tag{110-3}$$

The coefficients $\frac{f^{(n)}(x_0)}{n!}$, $(n = 1, 2, \cdots)$, appearing in (110-3) can be calculated from the known relationship (110-1) by the rule for differentiation of implicit functions.

The differentiation of (110-1) with respect to x gives

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0$$

and, if $F_y \neq 0$ at the point (x_0, y_0), then

$$\left.\frac{dy}{dx}\right|_{x=x_0} \equiv f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}.$$

The value of $f''(x)$ at $x = x_0$ can be calculated from the formula

$$\text{(110-4)} \qquad \frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)},$$

which implies that F_y does not vanish in some neighborhood of the point (x_0, y_0). Then

$$\text{(110-5)} \qquad \frac{d^2y}{dx^2} = -\frac{F_y\left(F_{xx} + F_{xy}\dfrac{dy}{dx}\right) - F_x\left(F_{yx} + F_{yy}\dfrac{dy}{dx}\right)}{F_y^2},$$

and making use of (110-4) leads to the formula

$$\text{(110-6)} \qquad \frac{d^2y}{dx^2} = -\frac{F_y^2F_{xx} - 2F_xF_yF_{xy} + F_x^2F_{yy}}{F_y^3},$$

provided that the inversion of the order of differentiation is legitimate.*

The formula (110-6) can be used to calculate the $f''(x_0)$ appearing in (110-3), since the values of the second partial derivatives at the point (x_0, y_0) can be calculated from the given function $F(x, y)$. The value of $f'''(x)$ at $x = x_0$ can be calculated with the aid of (110-6), and so on. In this manner one can construct the solution of (110-1) for y in the neighborhood of the point (x_0, y_0) in the form (110-3). The essential feature of this discussion is the requirement that $F_y(x, y)$ does not vanish in the vicinity of the point (x_0, y_0).

As an illustration of this approach to the problem of implicit functions consider the task of obtaining the solution of

$$x^2 + y^2 = 5,$$

---

* See Sec. 31.

which is valid in the vicinity of the point (2, 1). Noting that

$$2x + 2y\frac{dy}{dx} = 0$$

gives

(110-7) $$\frac{dy}{dx} = -\frac{x}{y}.$$

Differentiating again gives

(110-8) $$\frac{d^2y}{dx^2} = -\frac{y - y'x}{y^2} = -\frac{x^2 + y^2}{y^3},$$

where the last step is obtained by substituting the expression for y' from (110-7). The expression for $\frac{d^3y}{dx^3}$ can be calculated from (110-8); thus,

$$\frac{d^3y}{dx^3} = -\frac{2y^4y' + 2xy^3 - 3y^2y'(x^2 + y^2)}{y^6}$$
$$= -\frac{3x(x^2 + y^2)}{y^5}.$$

Obviously, the process of successive differentiation can be continued, in this case, as many times as desired. Substituting $x = 2$ and $y = 1$ in the formulas for the derivatives just obtained gives

$$f'(2) = -2, \qquad f''(2) = -5, \qquad f'''(2) = -30,$$

so that the solution (110-3) becomes

(110-9) $$y = 1 - 2(x - 2) - \frac{5}{2!}(x - 2)^2 - \frac{30}{3!}(x - 2)^3 + \cdots$$

Of course, the solution of

$$x^2 + y^2 = 5,$$

in the vicinity of the point (2, 1), could have been obtained by elementary algebra to give

$$y = \sqrt{5 - x^2},$$

which upon expansion in the series of powers of $x - 2$ will yield the infinite series (110-9).

111. Generalization of the Simple Problem. The discussion of Sec. 110 may be generalized in two directions:

(*a*) A greater number of variables may be introduced in the functional relationship (110-1);

(*b*) The number of functions may be increased.

As an instance of the generalization of the first type, consider a relationship of the form

$$F(x, y, z) = 0, \tag{111-1}$$

and assume that corresponding to a certain region of values of x and y in the neighborhood of the point (x_0, y_0) the equation (111-1) defines the function z, say

$$z = f(x, y). \tag{111-2}$$

It is understood that (111-2) is such a solution of (111-1) that

$$F(x_0, y_0, z_0) \equiv 0.$$

Assuming, as in the case of the simple problem treated in Sec. 110, that $f(x, y)$ can be expressed in the form of Taylor's series, namely,

$$z = f(x_0, y_0) + \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) + \cdots, \tag{111-3}$$

one can calculate the partial derivatives entering in (111-3) from (111-1). Thus,

$$dF = 0 = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy + \frac{\partial F}{\partial z}\,dz,$$

and since x and y are assumed to be independent variables, it follows that

$$\frac{\partial z}{\partial x} = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial z}}, \qquad \frac{\partial z}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}}, \tag{111-4}$$

unless $\dfrac{\partial F}{\partial z}$ vanishes in the neighborhood of (x_0, y_0, z_0).

The substitution of $x = x_0$, $y = y_0$, $z = z_0$ in (111-4) yields the coefficients of the linear terms appearing in (111-3). The coefficients of the higher powers of $x - x_0$ and $y - y_0$ can be calculated successively with the aid of the formulas (111-4). Again, it should be noted that it is necessary to demand that $\frac{\partial F}{\partial z}$ does not vanish in some neighborhood of the point (x_0, y_0, z_0), if the formulas (111-4) are to have a meaning.

As an example of the generalization of the second type, consider a pair of equations

$$\left\{\begin{aligned} F(x, y, u, v) &= 0, \\ G(x, y, u, v) &= 0, \end{aligned}\right. \tag{111-5}$$

which connect four variables x, y, u, and v. Let it be assumed that there exists a set of four real numbers x_0, y_0, u_0, v_0 such that

$$F(x_0, y_0, u_0, v_0) \equiv 0,$$

and

$$G(x_0, y_0, u_0, v_0) \equiv 0.$$

Further, suppose that the equations (111-5) can be solved in the vicinity of the point (x_0, y_0, u_0, v_0) to yield

$$\left\{\begin{aligned} u &= f(x, y), \\ v &= g(x, y), \end{aligned}\right. \tag{111-6}$$

which are so chosen that

$$\left\{\begin{aligned} u_0 &= f(x_0, y_0), \\ v_0 &= g(x_0, y_0). \end{aligned}\right. \tag{111-7}$$

If the functions $f(x, y)$ and $g(x, y)$ possess sufficiently good properties to admit of power-series expansion in the neighborhood of the point (x_0, y_0), then one can write

$$\left\{\begin{aligned} u &= f(x_0, y_0) + \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) \\ &\qquad + \cdots, \\ v &= g(x_0, y_0) + \left(\frac{\partial g}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial g}{\partial y}\right)_0 (y - y_0) \\ &\qquad + \cdots. \end{aligned}\right. \tag{111-8}$$

The values $f(x_0, y_0)$ and $g(x_0, y_0)$ are known, since the left-hand members of (111-7) are prescribed in advance. The coefficients of the linear and higher powers of $x - x_0$ and $y - y_0$ can be determined by the differentiation of equations (111-5), where u and v are regarded as functions of the independent variables x and y.

Thus, differentiating equations (111-5) with respect to x gives

$$\begin{aligned}\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} &= 0,\\ \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} &= 0.\end{aligned}$$

Solving this system for $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial v}{\partial x}$ by Cramer's rule provides

$$(111\text{-}9)\qquad \frac{\partial u}{\partial x} = -\frac{\begin{vmatrix}\dfrac{\partial F}{\partial x} & \dfrac{\partial F}{\partial v}\\ \dfrac{\partial G}{\partial x} & \dfrac{\partial G}{\partial v}\end{vmatrix}}{\begin{vmatrix}\dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v}\\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v}\end{vmatrix}}, \qquad \frac{\partial v}{\partial x} = -\frac{\begin{vmatrix}\dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial x}\\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial x}\end{vmatrix}}{\begin{vmatrix}\dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v}\\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v}\end{vmatrix}}.$$

Now, if the denominator

$$(111\text{-}10)\qquad J \equiv \begin{vmatrix}\dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v}\\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v}\end{vmatrix}$$

does not vanish for (x_0, y_0, u_0, v_0), the values of

$$\left(\frac{\partial u}{\partial x}\right)_0 \equiv \left(\frac{\partial f}{\partial x}\right)_0 \quad \text{and} \quad \left(\frac{\partial v}{\partial x}\right)_0 \equiv \left(\frac{\partial g}{\partial x}\right)_0$$

can be calculated from (111-9). Similarly, the differentiation of equations (111-5) with respect to y will yield formulas for $\dfrac{\partial u}{\partial y}$ and $\dfrac{\partial v}{\partial y}$ which are analogous to (111-9) and which contain the same determinant J in the denominator.

The coefficients of higher powers of $x - x_0$ and $y - y_0$ in (111-8) can be calculated with the aid of the expressions for the first partial derivatives. It is obvious that the success of this scheme depends on the nonvanishing of the determinant (111-10) in the neighborhood of the point (x_0, y_0, u_0, v_0).

An entirely similar consideration of the set of three equations

$$\text{(111-11)} \qquad \begin{cases} F(x, y, z, u, v, w) = 0, \\ G(x, y, z, u, v, w) = 0, \\ H(x, y, z, u, v, w) = 0, \end{cases}$$

shows that the determinant

$$\text{(111-12)} \qquad J \equiv \begin{vmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} & \dfrac{\partial F}{\partial w} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} & \dfrac{\partial G}{\partial w} \\ \dfrac{\partial H}{\partial u} & \dfrac{\partial H}{\partial v} & \dfrac{\partial H}{\partial w} \end{vmatrix}$$

plays the same role in the problem of solving (111-11) for u, v, and w in terms of x, y, and z as (111-10) did in the simpler case.

The determinants of the form (111-10) and (111-12) are called *Jacobians*, and they are fundamental to all of the investigations of this chapter.

A particular form of the functional relationship (111-5) is of frequent occurrence in geometry. Let

$$\text{(111-13)} \qquad \begin{cases} x = \varphi(u, v), \\ y = \psi(u, v), \end{cases}$$

be regarded as the equations of transformation which establish a relationship between a set of points of some region of the xy-plane with the corresponding set of points in the uv-plane. It will be assumed that there is a one-to-one correspondence of the points of the region in the xy-plane with those of the uv-plane. Let it be required to obtain from (111-13) the inverse transformation, namely,

$$\text{(111-14)} \qquad \begin{cases} u = f(x, y), \\ v = g(x, y). \end{cases}$$

The equations (111-13) can be put in the form (111-5) by writing them as

$$F(x, y, u, v) \equiv \varphi(u, v) - x = 0,$$
$$G(x, y, u, v) \equiv \psi(u, v) - y = 0.$$

Hence, if the scheme outlined above is to be available for the calculation of the inverse transformation, it is necessary to demand the nonvanishing of the determinant

$$(111\text{-}15)\qquad J = \begin{vmatrix} \dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\varphi}{\partial v} \\ \dfrac{\partial\psi}{\partial u} & \dfrac{\partial\psi}{\partial v} \end{vmatrix}.$$

It is clear that a similar consideration applied to a set of n equations of transformation of the form

$$(111\text{-}16)\qquad \begin{cases} u_1 = u_1(x_1, x_2, \cdots, x_n), \\ u_2 = u_2(x_1, x_2, \cdots, x_n), \\ \cdots\cdots\cdots\cdots\cdots, \\ u_n = u_n(x_1, x_2, \cdots, x_n), \end{cases}$$

will lead to a consideration of the determinant

$$(111\text{-}17)\qquad J \equiv \begin{vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}$$

It is interesting to note that if the equations of the transformation (111-16) are linear, so that

$$u_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n, \qquad (i = 1, 2, \cdots, n),$$

then (111-17) is precisely the determinant of the coefficients of the x's, namely,

$$\Delta \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

and it is well known that in obtaining the solution for the variables x_i in terms of the u_i, it is necessary to demand the nonvanishing of the determinant Δ.

112. Functional Dependence. Another important problem arising in the study of implicit functions is that of the functional relationship that may exist among a set of functions. As a simple example of such a circumstance, consider a pair of real functions

$$\begin{cases} u = f(x, y), \\ v = g(x, y), \end{cases} \tag{112-1}$$

of the variables x and y, defined in some region R of the xy-plane. It will be supposed that the functions u and v, together with their first partial derivatives, are continuous in R. Further, suppose that there exists a functional relationship between the functions u and v of the form

$$F(u, v) = 0, \tag{112-2}$$

such that

$$F[f(x, y), g(x, y)] \equiv 0. \tag{112-3}$$

The identity sign in (112-3) means that the left-hand member of (112-2) annuls itself for all values of x and y in the region R when $f(x, y)$ and $g(x, y)$ are substituted for u and v, respectively.

Differentiating (112-2) with respect to x and y gives

$$\begin{cases} \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial x} = 0, \\ \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial y} = 0. \end{cases} \tag{112-4}$$

The equations (112-4) may be looked upon as a set of two homogeneous linear equations in $\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial v}$. Now $\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial v}$ cannot both vanish identically in R, since this would imply that $F(u, v)$ is independent of u and v. Consequently, if F is to be a function of u and v, then the system (112-4) must possess nonzero solutions for $\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial v}$; hence, the determinant of the coefficients of $\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial v}$, namely

$$J = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} \end{vmatrix},$$

must vanish for all values of x and y in the region R. These considerations can obviously be extended to cover the case of a set of n functions $u_i(x_1, x_2, \cdots, x_n)$ of n independent variables $x_1, x_2, \cdots, x_n$.

It appears from this formal investigation that the vanishing of the Jacobian is a necessary condition for the existence of a functional relationship among a set of n functions $u_i(x_1, x_2, \cdots, x_n)$, and it will be shown rigorously in Sec. 115 that this condition is both necessary and sufficient.

As a specific illustration of the foregoing consider the set of three functions

$$\begin{cases} u = x + y + z, \\ v = xy + yz + xz, \\ w = x^2 + y^2 + z^2. \end{cases} \tag{112-5}$$

It is easily checked that

$$F(u, v, w) \equiv u^2 - 2v - w = 0.$$

The Jacobian of (112-5), namely,

$$\begin{vmatrix} 1 & y + z & 2x \\ 1 & x + z & 2y \\ 1 & y + x & 2z \end{vmatrix} \equiv 0$$

as it should be.

On the other hand, the relations

$$u = x^2 - y^2,$$
$$v = 2xy,$$

give

$$J = \begin{vmatrix} 2x & 2y \\ -2y & 2x \end{vmatrix} = 4(x^2 + y^2),$$

which obviously does not vanish identically in any region R, so that u and v cannot be connected by means of any analytic relationship of the form $F(u, v) = 0$.

113. Existence Theorem for Implicit Functions. It is the purpose of this section to furnish a rigorous discussion of the problem which was treated formally in Sec. 110. Let $F(x, y)$, regarded as a function of two independent variables x and y, be such that for $x = x_0$ and $y = y_0$

$$F(x_0, y_0) = 0.$$

Consider a rectangular region (Fig. 89) bounded by the lines

$$x = x_0 \pm h,$$
$$y = y_0 \pm k,$$

and suppose that corresponding to every value of x in the interval $(x_0 - h, x_0 + h)$ there exists one, and only one, value of y lying in the interval $(y_0 - k, y_0 + k)$ such that

$$F(x, y) = 0. \tag{113-1}$$

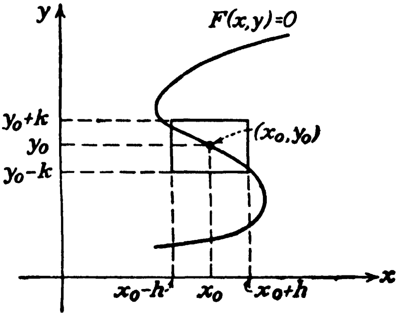

FIG. 89.

Under these circumstances the equation (113-1) is said to define y as an implicit function of x. If this function be denoted by the symbol

$$y = f(x),$$

then

$$y_0 = f(x_0),$$

and the latter is called the solution of (113-1) for y in terms of x at the point (x_0, y_0).

The following existence theorem is of fundamental importance in all considerations of this chapter.

Theorem 1. *Let x_0 be any real value of x such that:*

(*a*) $F(x_0, y) = 0$ *has a real solution* y_0, *that is,*

$$F(x_0, y_0) = 0;$$

(*b*) $F(x, y)$, *regarded as a function of two independent variables* x *and* y, *is continuous and has continuous partial derivatives* $F_x(x, y)$ *and* $F_y(x, y)$ *in some region* R *enclosing the point* (x_0, y_0);

(c) $F_y(x_0, y_0) \neq 0$;
then $F(x, y) = 0$ *can be solved (theoretically) as* $y = f(x)$ *in the vicinity of* $x = x_0$ *in such a way that*

$$y_0 = f(x_0).$$

Moreover, the solution is unique and is such that $y = f(x)$ *has a continuous derivative in the vicinity of* $x = x_0$ *which is given by the formula*

$$f'(x) = -\frac{F_x(x, y)}{F_y(x, y)}.$$

Since $F(x, y)$ is assumed to be continuous in some region containing the point (x_0, y_0), the function $F(x_0, y)$, obtained from $F(x, y)$ by setting $x = x_0$, is a continuous* function of y. Furthermore, $\frac{\partial F}{\partial y} \neq 0$ at (x_0, y_0), so that the function $F(x_0, y)$ is increasing in the vicinity of $y = y_0$ if $F_y(x_0, y_0) > 0$, and is decreasing if $F_y(x_0, y_0) < 0$. But $F(x_0, y_0) = 0$; hence, the function $F(x_0, y)$ must change sign at $y = y_0$. For definiteness let $F_y(x_0, y_0) > 0$; then there exists an interval $(y_0 - \delta, y_0 + \delta)$ such that the function $F(x_0, y)$ is negative for all values of y such that

$$y_0 - \delta \leqq y < y_0,$$

and is positive whenever

$$y_0 < y \leqq y_0 + \delta.$$

Furthermore, $F(x, y_0 - \delta)$ and $F(x, y_0 + \delta)$ are continuous functions of x, and for $x = x_0$ the function $F(x, y_0 - \delta)$ is negative. Consequently, there exists some interval $(x_0 - \delta', x_0 + \delta')$ about the point $x = x_0$ such that $F(x, y_0 - \delta) < 0$ for all values of x in this interval.

Similarly, $F(x_0, y_0 + \delta) > 0$, so that in a sufficiently small interval about the point $x = x_0$, $F(x, y_0 + \delta)$ will be positive. Let this interval be $(x_0 - \delta'', x_0 + \delta'')$, and if δ_1 is the smaller of the two numbers δ' and δ'', the functions $F(x, y_0 - \delta)$ and $F(x, y_0 + \delta)$ will have opposite signs for all values of x that are interior to the interval $(x_0 - \delta_1, x_0 + \delta_1)$ (Fig. 90),

---

* See Sec. 22.

Let $x = x_1$ be any value of x in the interval $(x_0 - \delta_1, x_0 + \delta_1)$; then, since $F(x_1, y_0 - \delta)$ and $F(x_1, y_0 + \delta)$ are of opposite signs, there will be some value of y, say $y = y_1$, where

$$y_0 - \delta < y_1 < y_0 + \delta,$$

such that

$$F(x_1, y_1) = 0. \tag{113-2}$$

The value of y satisfying (113-2) depends upon the choice of x, and, since for each value of x that is interior to $(x_0 - \delta_1, x_0 + \delta_1)$ there is determined a value of y satisfying $F(x, y) = 0$, one can write

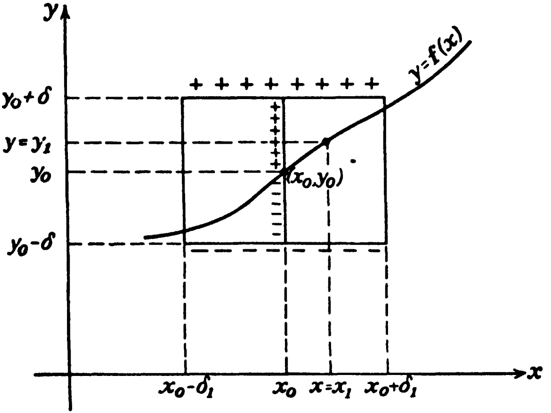

FIG. 90.

$$y = f(x).$$

The continuity of $y = f(x)$ in the vicinity of $x = x_0$ follows from the fact that for all values of x in the interval $(x_0 - \delta_1, x_0 + \delta_1)$ the equation

$$F(x, y) = 0$$

has a solution $y = f(x)$, for which y lies between $y_0 - \delta$ and $y_0 + \delta$.

At this stage of the proof one cannot assert that there is but one value of y corresponding to the choice $x = x_1$. In order to establish the uniqueness of the solution, the hypothesis of the continuity of $F_y(x, y)$ in some region enclosing the point

(x_0, y_0) will be utilized. If $F_y(x, y)$ is continuous in the vicinity of (x_0, y_0) and is different from zero at (x_0, y_0), then it will be different from zero at all points sufficiently near (x_0, y_0). Now, if in the foregoing discussion δ and δ_1 are chosen so small that

$$F_y(x, y) \neq 0 \qquad \text{when} \qquad |x - x_0| < \delta_1 \qquad \text{and} \qquad |y - y_0| < \delta,$$

then the function

$$F(x_1, y), \qquad \text{where} \qquad x_0 - \delta_1 < x_1 < x_0 + \delta_1,$$

will be either increasing or decreasing for all values of y in the interval $(y_0 - \delta, y_0 + \delta)$. Under these circumstances $F(x_1, y)$ cannot vanish for more than one value of y.

In order to establish the continuity of the derivative of $y = f(x)$ in the vicinity of $x = x_0$, let $(x + \Delta x, y + \Delta y)$ and (x, y) be a pair of solutions of $F(x, y) = 0$. Then

$$F(x + \Delta x, y + \Delta y) - F(x, y) = 0$$

and, obviously,

$$F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) + F(x + \Delta x, y) - F(x, y) = 0.$$

By hypothesis $F_x(x, y)$ and $F_y(x, y)$ are continuous functions in some region R about (x_0, y_0). Therefore, from the mean-value theorem,* it follows that

$$\Delta y F_y(x + \Delta x, y + \theta_1 \Delta y) + \Delta x F_x(x + \theta_2 \Delta x, y) = 0,$$

where $0 < \theta_1 < 1$ and $0 < \theta_2 < 1$. Solving for $\frac{\Delta y}{\Delta x}$ gives

$$\frac{\Delta y}{\Delta x} = -\frac{F_x(x + \theta_2 \Delta x, y)}{F_y(x + \Delta x, y + \theta_1 \Delta y)}$$

and, letting $\Delta x \to 0$, one obtains

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)},$$

since $F_y(x, y) \neq 0$ for every point (x, y) in the region R.

If the assumption of the continuity of $F_x(x, y)$ is abrogated, one cannot assert that the solution $y = f(x)$ will possess a continuous derivative in the vicinity of $x = x_0$. However, the solution $y = f(x)$ will be unique and continuous under the hypothesis that $F_y(x, y)$ is continuous in the neighborhood of (x_0, y_0) and is different from zero at the point (x_0, y_0). It is not difficult to demonstrate the existence of a unique solution on the hypotheses that $F_y(x_0, y_0)$ is different from zero and $F_y(x, y)$ does not change sign in the vicinity of (x_0, y_0).*

---

* See Sec. 20.

An argument in every respect similar to the foregoing permits one to generalize Theorem 1 to include the case of functions of more than two variables. Thus, one can enunciate the following theorem:

Theorem 2. *Let the function* $F(x_1, x_2, \cdots, x_m, u)$ *satisfy the following conditions:*

(*a*) $F(x_1, x_2, \cdots, x_m, u)$ *has a real solution* u^0 *for* $x_1 = x_1^0$, $x_2 = x_2^0, \cdots, x_m = x_m^0$;

(*b*) $F(x_1, x_2, \cdots, x_m, u)$, *regarded as a function of the* $m + 1$ *independent variables* $x_1, x_2, \cdots, x_m, u$, *is continuous and has continuous partial derivatives* $F_{x_1}, F_{x_2}, \cdots, F_{x_m}, F_u$ *in some region of space enclosing the point* $(x_1^0, x_2^0, \cdots, x_m^0, u^0)$;

(*c*) $F_u(x_1^0, x_2^0, \cdots, x_m^0, u^0) \neq 0$.

Then $F(x_1, x_2, \cdots, x_m, u) = 0$ *can be solved in the vicinity of* $(x_1^0, x_2^0, \cdots, x_m^0)$ *to yield*

$$u = f(x_1, x_2, \cdots, x_m)$$

in such a way that

$$u^0 = f(x_1^0, x_2^0, \cdots, x_m^0).$$

The solution $u = f(x_1, x_2, \cdots, x_m)$ *is unique and is continuous in the vicinity of the point* $(x_1^0, x_2^0, \cdots, x_m^0)$ *together with its first partial derivatives, which are given by the formulas*

$$\begin{aligned}
\frac{\partial u}{\partial x_1} &= -\frac{F_{x_1}(x_1, x_2, \cdots, x_m, u)}{F_u(x_1, x_2, \cdots, x_m, u)}, \\
\frac{\partial u}{\partial x_2} &= -\frac{F_{x_2}(x_1, x_2, \cdots, x_m, u)}{F_u(x_1, x_2, \cdots, x_m, u)}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots, \\
\frac{\partial u}{\partial x_m} &= -\frac{F_{x_m}(x_1, x_2, \cdots, x_m, u)}{F_u(x_1, x_2, \cdots, x_m, u)}.
\end{aligned}$$

---

* See Hobson, E. W., Functions of a Real Variable, vol. 1, Secs. 316, 316^1, pp. 432–436.

114. Existence Theorem for Simultaneous Equations.

Theorem. *Let a system of n equations*

$$(114\text{-}1)\quad F_i(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_n) = 0, \qquad (i = 1, 2, \cdots, n),$$

in $m + n$ variables $x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_n$ have a real solution $x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_n^0$, so that

$$(114\text{-}2)\quad F_i(x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_n^0) = 0, \qquad (i = 1, 2, \cdots, n).$$

If the functions

$$F_i(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_n),$$

regarded as functions of $m + n$ independent variables, are continuous and have continuous first partial derivatives in some region R of the $m + n$ dimensional space enclosing the point P, whose coordinates are $(x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_n^0)$, and if the Jacobian

$$J \equiv \frac{\partial(F_1, F_2, \cdots, F_n)}{\partial(u_1, u_2, \cdots, u_n)} \equiv \begin{vmatrix} \frac{\partial F_1}{\partial u_1} & \frac{\partial F_1}{\partial u_2} & \cdots & \frac{\partial F_1}{\partial u_n} \\ \frac{\partial F_2}{\partial u_1} & \frac{\partial F_2}{\partial u_2} & \cdots & \frac{\partial F_2}{\partial u_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_n}{\partial u_1} & \frac{\partial F_n}{\partial u_2} & \cdots & \frac{\partial F_n}{\partial u_n} \end{vmatrix}$$

is different from zero at the point P, then the set of equations (114-1) *can be solved as*

$$(114\text{-}3)\quad u_i = f_i(x_1, x_2, \cdots, x_m), \qquad (i = 1, 2, \cdots, n),$$

in the vicinity of the point P in such a way that

$$u_i^0 = f_i(x_1^0, x_2^0, \cdots, x_m^0), \qquad (i = 1, 2, \cdots, n).$$

Moreover, the set of solutions (114-3) *is unique and the functions u_i are continuous together with their first partial derivatives in the neighborhood of the point $(x_1^0, x_2^0, \cdots, x_m^0)$.*

This theorem has been established in Theorem 2, Sec. 113, for the case when $n = 1$, in which the Jacobian reduces to F_{u_1}.

It is natural to extend this proof to a greater number of functions by mathematical induction. Thus, if the correctness of the theorem can be established for n functions on the hypothesis that the theorem is true for $n - 1$ functions, then the truth of the theorem follows for $n = 2, 3, 4, \cdots$, since the theorem is known to hold for $n = 1$.

By hypothesis the determinant J does not vanish for the particular set of values $(x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_n^0)$, so that all of its elements cannot be equal to zero. Let it be supposed that it is the partial derivative

$$\left.\frac{\partial F_n}{\partial u_n}\right|_P,$$

appearing in J, that does not vanish.* Then it follows from Theorem 2, Sec. 113, that the equation

$$F_n(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_n) = 0 \tag{114-4}$$

can be solved to yield

$$u_n = \varphi(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_{n-1}), \tag{114-5}$$

in such a way that

$$u_n^0 = \varphi(x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_{n-1}^0).$$

Substituting (114-5) in the first $n - 1$ equations of the set (114-1) gives

$$\Phi_i(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_{n-1}) = 0, \qquad (i = 1, 2, \cdots, n - 1), \tag{114-6}$$

where

$$\Phi_i(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_{n-1}) \equiv F_i(x_1, x_2, \cdots, x_m, u_1, u_2, \cdots, u_{n-1}, \varphi). \tag{114-7}$$

Equations (114-5) and (114-6) are equivalent to the system (114-1). Now if the $n - 1$ equations (114-6) can be solved for $u_1, u_2, \cdots, u_{n-1}$ in terms of $x_1, x_2, \cdots, x_m$, there will be obtained a set of $n - 1$ functions

$$u_i = f_i(x_1, x_2, \cdots, x_m), \qquad (i = 1, 2, \cdots, n - 1). \tag{114-8}$$

---

* By reordering the functions (114-1), if necessary, it is always possible to achieve this.

Substituting from (114-8) in (114-5) gives

$$u_n = f_n(x_1, x_2, \cdots, x_m). \tag{114-9}$$

The set of n functions (114-8) and (114-9) is the desired set (114-3). But the set of $n - 1$ equations (114-6) is of the type (114-1), and if it is assumed that the theorem is valid for $n - 1$ functions, then Eqs. (114-6) can be solved for $u_1, u_2, \cdots, u_{n-1}$ in terms of $x_1, x_2, \cdots, x_m$, provided that

$$\Delta \equiv \frac{\partial(\Phi_1, \Phi_2, \cdots, \Phi_{n-1})}{\partial(u_1, u_2, \cdots, u_{n-1})} \neq 0$$

for $(x_1^0, x_2^0, \cdots, x_m^0, u_1^0, u_2^0, \cdots, u_{n-1}^0)$.

In order to show that Δ does not vanish it is desirable to write it out explicitly. From (114-7) it follows that

$$\frac{\partial \Phi_i}{\partial u_j} = \frac{\partial F_i}{\partial u_j} + \frac{\partial F_i}{\partial u_n}\frac{\partial \varphi}{\partial u_j}, \qquad (i, j = 1, 2, \cdots, n - 1),$$

so that

$$\Delta = \begin{vmatrix} \dfrac{\partial F_1}{\partial u_1} + \dfrac{\partial F_1}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_1} & \dfrac{\partial F_1}{\partial u_2} + \dfrac{\partial F_1}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_2} & \cdots & \dfrac{\partial F_1}{\partial u_{n-1}} + \dfrac{\partial F_1}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_{n-1}} \\ \dfrac{\partial F_2}{\partial u_1} + \dfrac{\partial F_2}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_1} & \dfrac{\partial F_2}{\partial u_2} + \dfrac{\partial F_2}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_2} & \cdots & \dfrac{\partial F_2}{\partial u_{n-1}} + \dfrac{\partial F_2}{\partial u_n} \cdot \dfrac{\partial \varphi}{\partial u_{n-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial F_{n-1}}{\partial u_1} + \dfrac{\partial F_{n-1}}{\partial u_n} \dfrac{\partial \varphi}{\partial u_1} & \dfrac{\partial F_{n-1}}{\partial u_2} + \dfrac{\partial F_{n-1}}{\partial u_n} \dfrac{\partial \varphi}{\partial u_2} & \cdots & \dfrac{\partial F_{n-1}}{\partial u_{n-1}} + \dfrac{\partial F_{n-1}}{\partial u_n} \dfrac{\partial \varphi}{\partial u_{n-1}} \end{vmatrix}.$$

Since each element of this determinant is made up of the sum of two terms, Δ can be expressed as the sum of 2^{n-1} determinants.* Some of these 2^{n-1} determinants will contain two columns that are proportional, so that they will vanish, and the remaining ones give

$$\Delta = \frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_1, u_2, \cdots, u_{n-1})} + \frac{\partial \varphi}{\partial u_1}\frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_n, u_2, \cdots, u_{n-1})} + \cdots + \frac{\partial \varphi}{\partial u_{n-1}}\frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_1, u_2, \cdots, u_{n-2}, u_n)}. \tag{114-10}$$

---

* It will be recalled that

$$\begin{vmatrix} a_1 + b_1 & c_1 & d_1 \\ a_2 + b_2 & c_2 & d_2 \\ a_3 + b_3 & c_3 & d_3 \end{vmatrix} = \begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix} + \begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix}.$$

The expressions for

$$\frac{\partial \varphi}{\partial u_1}, \frac{\partial \varphi}{\partial u_2}, \cdots, \frac{\partial \varphi}{\partial u_{n-1}}$$

which appear in (114-10), can be obtained as follows. Recalling (114-5) and applying the rule for the differentiation of implicit functions to (114-4), one finds

$$\frac{\partial F_n}{\partial u_i} + \frac{\partial F_n}{\partial u_n}\frac{\partial \varphi}{\partial u_i} = 0, \qquad (i = 1, 2, \cdots, n-1).$$

Hence,

$$\frac{\partial \varphi}{\partial u_i} = -\frac{\dfrac{\partial F_n}{\partial u_i}}{\dfrac{\partial F_n}{\partial u_n}}.$$

On substituting these expressions for the $\frac{\partial \varphi}{\partial u_i}$ in (114-10) and multiplying both sides of the resulting equation by $\frac{\partial F_n}{\partial u_n}$, there results

$$\begin{aligned}\Delta \cdot \frac{\partial F_n}{\partial u_n} = \frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_1, u_2, \cdots, u_{n-1})} \cdot \frac{\partial F_n}{\partial u_n} \\ - \frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_n, u_2, \cdots, u_{n-1})} \cdot \frac{\partial F_n}{\partial u_1} - \cdots \\ - \frac{\partial(F_1, F_2, \cdots, F_{n-1})}{\partial(u_1, u_2, \cdots, u_{n-2}, u_n)} \cdot \frac{\partial F_n}{\partial u_{n-1}}.\end{aligned}$$

A little reflection shows that the right-hand member of this equation is precisely the expansion of the Jacobian J in terms of the elements of the last row of J. Hence,

$$\Delta = \frac{J}{\dfrac{\partial F_n}{\partial u_n}},$$

which, by hypothesis, does not vanish at the point P.

115. Functional Dependence.

Theorem. *Let* $u_i = f_i(x_1, x_2, \cdots, x_n)$, $(i = 1, 2, \cdots, n)$, *be a set of n functions of n independent variables $x_1, x_2, \cdots, x_n$. The functions f_i together with their first partial derivatives are* *assumed to be continuous in some region R enclosing the point* $(x_1^0, x_2^0, \cdots, x_n^0)$. *A necessary and sufficient condition that there exist among these functions a relation*

$$F(u_1, u_2, \cdots, u_n) = 0,$$

which does not contain the variables $x_1, x_2, \cdots, x_n$ *explicitly, is that*

$$J = \frac{\partial(u_1, u_2, \cdots, u_n)}{\partial(x_1, x_2, \cdots, x_n)} \equiv 0$$

for all values of the variables $x_1, x_2, \cdots, x_n$ *in the region R.*

In order to avoid complications in writing, the theorem will be established for the case of three functions

$$\begin{cases} u = f_1(x, y, z), \\ v = f_2(x, y, z), \\ w = f_3(x, y, z), \end{cases} \tag{115-1}$$

since the discussion of this case possesses all features of the more general circumstance.

To demonstrate the necessity of the condition of the theorem, assume that the Jacobian

$$J = \frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z} \end{vmatrix} \tag{115-2}$$

does not vanish at some point (x_0, y_0, z_0) of the region R. Denote the values of u, v, and w, calculated from (115-1) by setting $x = x_0$, $y = y_0$, $z = z_0$, by u_0, v_0, and w_0. It follows from the theorem of Sec. 114 that one can solve (115-1) to give

$$\begin{aligned} x &= \varphi_1(u, v, w), \\ y &= \varphi_2(u, v, w), \\ z &= \varphi_3(u, v, w), \end{aligned}$$

where the solution is valid for an arbitrary choice of u, v, and w belonging to some region about the point (u_0, v_0, w_0). Since the choice of the numbers u, v, w is arbitrary, there can be no relation of the form

$$F(u, v, w) = 0$$

in the region R under consideration.

To prove the sufficiency, consider again the set of functions (115-1) and assume that (115-2) vanishes identically in the region R. Consider first the case where at least one of the minors of J does not vanish identically, and assume that such a one is

$$\Delta = \begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{vmatrix}$$

Since $\Delta \not\equiv 0$ in R, one can solve the first two of Eqs. (115-1) for x and y to yield

$$\text{(115-3)} \qquad \begin{cases} x = \varphi_1(u, v, z), \\ y = \varphi_2(u, v, z). \end{cases}$$

Substituting these values in the third of Eqs. (115-1) gives

$$\text{(115-4)} \qquad w = f_3(\varphi_1, \varphi_2, z) \equiv F(u, v, z).$$

Now if the right-hand member of (115-4) contains no z, then w is a function of the variables u and v and this will establish the sufficiency of the condition for this case. It will be shown next that $F(u, v, z)$ indeed is independent of z, since it will be proved that

$$\frac{\partial F}{\partial z} \equiv 0.$$

Calculating this partial derivative from (115-4) furnishes

$$\text{(115-5)} \qquad \frac{\partial F}{\partial z} = \frac{\partial f_3}{\partial x}\frac{\partial \varphi_1}{\partial z} + \frac{\partial f_3}{\partial y}\frac{\partial \varphi_2}{\partial z} + \frac{\partial f_3}{\partial z}.$$

But, since (115-3) represents the solution of the first two of Eqs. (115-1), the substitution of (115-3) in the first two of Eqs. (115-1) yields the identities

$$f_1(\varphi_1, \varphi_2, z) \equiv u,$$
$$f_2(\varphi_1, \varphi_2, z) \equiv v,$$

It is, therefore, evident that $f_1(\varphi_1, \varphi_2, z)$ and $f_2(\varphi_1, \varphi_2, z)$ are independent of z, so that

$$(115\text{-}6)\qquad \begin{cases} \dfrac{\partial f_1}{\partial x}\dfrac{\partial \varphi_1}{\partial z} + \dfrac{\partial f_1}{\partial y}\dfrac{\partial \varphi_2}{\partial z} + \dfrac{\partial f_1}{\partial z} = 0, \\ \dfrac{\partial f_2}{\partial x}\dfrac{\partial \varphi_1}{\partial z} + \dfrac{\partial f_2}{\partial y}\dfrac{\partial \varphi_2}{\partial z} + \dfrac{\partial f_2}{\partial z} = 0. \end{cases}$$

Adding to the last column of the determinant J the product of the first column by $\dfrac{\partial \varphi_1}{\partial z}$ and the product of the second column by $\dfrac{\partial \varphi_2}{\partial z}$ gives the determinant

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial x}\dfrac{\partial \varphi_1}{\partial z} + \dfrac{\partial f_1}{\partial y}\dfrac{\partial \varphi_2}{\partial z} + \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial x}\dfrac{\partial \varphi_1}{\partial z} + \dfrac{\partial f_2}{\partial y}\dfrac{\partial \varphi_2}{\partial z} + \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial x}\dfrac{\partial \varphi_1}{\partial z} + \dfrac{\partial f_3}{\partial y}\dfrac{\partial \varphi_2}{\partial z} + \dfrac{\partial f_3}{\partial z} \end{vmatrix},$$

which, upon noting (115-5) and (115-6), reduces to

$$J = \Delta \cdot \frac{\partial F}{\partial z}.$$

By hypothesis $J \equiv 0$ and $\Delta \not\equiv 0$, so that $\dfrac{\partial F}{\partial z} \equiv 0$.

It remains to consider the case where every minor of second order of the determinant J vanishes but where at least one of the elements of J does not vanish identically in R. Suppose, for example, that

$$\frac{\partial f_1}{\partial x} \not\equiv 0.$$

Then one can solve the first of Eqs. (115-1) for x in terms of y and z, so that

$$x = \varphi_1(u, y, z).$$

Substituting this value of x in the last two of Eqs. (115-1) gives

$$\begin{aligned} v &= f_2(\varphi_1, y, z) \equiv F(u, y, z), \\ w &= f_3(\varphi_1, y, z) \equiv G(u, y, z), \end{aligned}$$

and it remains to show that F and G do not contain y and z, in which event both v and w are functions of u. It will be proved next that such is the case.

Now,

$$\frac{\partial F}{\partial y} = \frac{\partial f_2}{\partial x}\frac{\partial \varphi_1}{\partial y} + \frac{\partial f_2}{\partial y},$$

and from the first of Eqs. (115-1) it follows that

$$\frac{\partial f_1}{\partial x}\frac{\partial \varphi_1}{\partial y} + \frac{\partial f_1}{\partial y} = 0.$$

From these equations, and from the fact that every minor of J of order 2 vanishes, it follows that

$$\frac{\partial F}{\partial y} \equiv 0.$$

It can be verified in a similar way that

$$\frac{\partial F}{\partial z} \equiv 0, \qquad \frac{\partial G}{\partial y} \equiv 0, \qquad \text{and} \qquad \frac{\partial G}{\partial z} \equiv 0,$$

and hence, both v and w are functions of u.

The remaining case in which every partial derivative vanishes identically is trivial, because in this event the functions u, v, and w reduce to mere constants.

The extension of this proof to any number of functions is obvious. It may be remarked that, if the Jacobian of the n functions

$$u_i = f_i(x_1, x_2, \cdots, x_n), \qquad (i = 1, 2, \cdots, n),$$

is such that every minor of order greater than r vanishes identically while at least one minor of order r does not vanish, then there will be exactly $n - r$ independent functional relations connecting the variables.

Example. Let

$$\begin{aligned} u &= x + 2y + z, \\ v &= x - 2y + z, \\ w &= x^2 + 2xz - 4y^2 + z^2 \end{aligned}$$

Here,

$$J = \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 2(x+z) & -8y & 2(x+z) \end{vmatrix} = 0,$$

so that the functions are dependent. Furthermore, since there exists a nonvanishing minor of second order, there is only one independent relation. In fact it is $w = uv$.

116. Properties of Jacobians. Let there be a set of n functions $F_1, F_2, \cdots, F_n$ of n variables $u_1, u_2, \cdots, u_n$, and suppose that the variables u_i are not independent variables but are functions of some other set of variables $x_1, x_2, \cdots, x_n$. Then it is easy to establish the following interesting formula:

$$(116\text{-}1)\qquad \frac{\partial(F_1, F_2, \cdots, F_n)}{\partial(x_1, x_2, \cdots, x_n)} = \frac{\partial(F_1, F_2, \cdots, F_n)}{\partial(u_1, u_2, \cdots, u_n)} \cdot \frac{\partial(u_1, u_2, \cdots, u_n)}{\partial(x_1, x_2, \cdots, x_n)}.$$

The result (116-1) follows upon multiplying the two determinants in the right-hand member according to the rule for the multiplication of determinants.* Thus,

$$\begin{vmatrix} \dfrac{\partial F_1}{\partial u_1} & \dfrac{\partial F_1}{\partial u_2} & \cdots & \dfrac{\partial F_1}{\partial u_n} \\ \dfrac{\partial F_2}{\partial u_1} & \dfrac{\partial F_2}{\partial u_2} & \cdots & \dfrac{\partial F_2}{\partial u_n} \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\ \dfrac{\partial F_n}{\partial u_1} & \dfrac{\partial F_n}{\partial u_2} & \cdots & \dfrac{\partial F_n}{\partial u_n} \end{vmatrix} \cdot \begin{vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}$$

---

* Note that if the n-rowed determinant $|a_{ij}|$, in which a_{ij} is the element in the ith row and the jth column, is multiplied by the n-rowed determinant $|b_{ij}|$, there results $|a_{ij}| \cdot |b_{ij}| = \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|$.

gives the determinant whose element in the ith row and the jth column is

$$\frac{\partial F_i}{\partial u_1}\frac{\partial u_1}{\partial x_j} + \frac{\partial F_i}{\partial u_2}\frac{\partial u_2}{\partial x_j} + \cdots + \frac{\partial F_i}{\partial u_n}\frac{\partial u_n}{\partial x_j}.$$

The latter is recognized as the formula for $\frac{\partial F_i}{\partial x_j}$. For $n = 1$, the formula (116-1) reduces to the simple formula for the derivative of a composite function, namely,

$$\frac{dF}{dx} = \frac{dF}{du}\cdot\frac{du}{dx}.$$

Another interesting property of Jacobians is worth noting. Consider a transformation

$$(116\text{-}2) \qquad u_i = f_i(x_1, x_2, \cdots, x_n), \qquad (i = 1, 2, \cdots, n),$$

and denote its inverse by

$$(116\text{-}3) \qquad x_i = \varphi_i(u_1, u_2, \cdots, u_n).$$

The Jacobians of (116-2) and (116-3) are

$$J_1 = \frac{\partial(u_1, u_2, \cdots, u_n)}{\partial(x_1, x_2, \cdots, x_n)} \quad \text{and} \quad J_2 = \frac{\partial(x_1, x_2, \cdots, x_n)}{\partial(u_1, u_2, \cdots, u_n)}.$$

Then,

$$J_1 \cdot J_2 = \frac{\partial(u_1, u_2, \cdots, u_n)}{\partial(x_1, x_2, \cdots, x_n)} \cdot \frac{\partial(x_1, x_2, \cdots, x_n)}{\partial(u_1, u_2, \cdots, u_n)},$$

which, by the formula (116-1), is equal to

$$\frac{\partial(u_1, u_2, \cdots, u_n)}{\partial(u_1, u_2, \cdots, u_n)} \equiv 1.$$

Hence, the Jacobians of the direct and the inverse transformations are reciprocals of one another.

PROBLEMS

1. If

$$\begin{aligned} u &= u(x, y, z), \\ v &= v(x, y, z), \\ w &= w(x, y, z), \end{aligned}$$

assume that $\dfrac{\partial(v, w)}{\partial(y, z)} \neq 0$, so that one can solve the last two equations for y and z. Substitute these solutions in the first equation, and show that

$$\frac{\partial u}{\partial x} = \frac{\partial(u, v, w)}{\partial(x, y, z)} \cdot \frac{\partial(y, z)}{\partial(v, w)}.$$

2. Let

$$\begin{aligned} x &= x(u, v), \\ y &= y(u, v), \\ z &= z(u, v), \end{aligned}$$

be the parametric equations of a surface from which u and v can be eliminated to give $z = F(x, y)$. Show that

$$\frac{\partial z}{\partial x} = \frac{\partial(y, z)}{\partial(u, v)} \div \frac{\partial(x, y)}{\partial(u, v)}.$$

3. The transformation of inversion in space is given by

$$u = \frac{x}{x^2 + y^2 + z^2}, \qquad v = \frac{y}{x^2 + y^2 + z^2}, \qquad w = \frac{z}{x^2 + y^2 + z^2}.$$

Does it possess an inverse? Verify that the product of the Jacobians of the direct and the inverse transformations is unity.

INDEX

A

Abel, N. H., 244
Abel's test for uniform convergence, 264
Abel's theorem, on interval of convergence, 268
  on power series, 276
Absolute convergence, of integrals, 337, 341, 344, 349
  of series, 236, 240, 263, 268
Addition of series, 212, 240, 281
Adiabatic process, 208
Alternating series, 234
Angle, as a line integral, 171
  solid, 171
Arc length, 127
Area, as a double integral, 140
  element of, 153, 201
  as a line integral, 178
  of surface, 161
Associative law for series, 212*n.*
Attraction, law of, 172, 202, 369, 371

B

Bernoulli's inequality, 8
Bernoulli's numbers, 287, 289
Beta function, 376
Binomial series, 301
Binomial theorem, 301
Bolzano-Weierstrass theorem, 18
Bounded sets, 105
Bounds, 22
  of a function, 23, 106

C

Cantor-Dedekind axiom, 3
Cauchy-Riemann equations, 206
Cauchy's integral test, 225
Cauchy's principal value, 347
Cauchy's product of series, 244
Cauchy's ratio test, 217*n.*
Cauchy's root test, 216
Cauchy's theorem, 53
Center of gravity, 137, 159, 165
Change of variables, in derivatives, 91
  in integrals, 124, 147, 155, 199
Circular functions, 308
Closed curve, area of, 178
  simple, 179
Comparison test for series, 215
Complex number, 3
Conditionally convergent series, 236
Conservative systems, 204
Continuity, 31, 35, 39, 58, 256, 276
  equation of, 206
  piece-wise, 34
  uniform, 37, 38, 60
Convergence, 11
  absolute, 236, 240, 263, 268, 337, 341, 344
  criterion, 13, 14, 20, 210, 233
  of integrals, 337, 341
    tests for, 341, 349, 367
  interval of, 248, 267, 269
  radius of, 269
  of series, 209, 247
    tests for, 215, 225, 237, 262
  uniform, 247, 263, 264, 355, 367, 408
Coordinate surfaces, 155
Coordinates, curvilinear, 148
  cylindrical, 157
  polar, 148, 152, 157
  spherical, 157
Curve, simple, 179
Curvilinear coordinates, 148
Cut, 2
Cylindrical coordinates, 157

D

D'Alembert's ratio test, 217
Darboux theorem, 108
Dedekind, 2, 3
Definite integrals, 99–129
  applications of, 126
  change of variables in, 124
  differentiation under sign of, 121
  evaluation of, 110, 118, 123, 144, 178, 309
  mean-value theorems on, 113
Density, property of, 1
Dependence, functional, 423, 433
Derivatives, 41, 47
  of composite functions, 45, 67, 71, 439
  directional, 76, 86, 87
  of higher order, 87
  of implicit functions, 68, 71, 89, 426
  of integrals, 119, 121
  normal, 78, 86, 87
  partial, 62
  of series, 261, 405
Differentials, 43
  exact, 191, 199
    (*See also* Total differential)
  of higher order, 47
  total, 64
Differentiation, of composite functions, 45, 67, 71, 439
  of implicit functions, 68, 71
    (*See also* Implicit functions)
  under integral sign, 121, 353, 367, 371
  inversion of the order, 87
  of series, 261, 405
Direction angles, 80
Direction components, 80
Direction cosines, 81, 162
Directional derivative, 76, 86
    (*See also* Gradient)
Dirichlet's conditions, 384, 385
Dirichlet's integral, 387, 362
Discontinuities, 34, 392
Divergence theorem, 167
Double integrals, 130
  improper, 365
Double series, 246
  convergence of, 247

E

e, base of natural logarithms, 28
Ellipse, length of, 314
Elliptic integrals, 117, 129, 312, 313, 352
Errors, approximate, 66
  relative, 66
Euler's formulas, 306
Euler's theorem, 75
Evaluation of integrals, 110, 118, 123, 131, 144, 178, 310, 357
Even function, 391
Exact differentials, 191, 199
Existence theorem, for implicit functions, 425
  for simultaneous equations, 430
Expansion, in Maclaurin's series, 298
  in power series, 290, 291, 298
  in series of orthogonal functions, 389
  in Taylor's series, 298
  in trigonometric series, 389
    (*See also* Fourier series)
Extended law of the mean, 291
Extremal values, 327, 332

F

Factorial, generalization of (*see* Gamma functions)
Fermat's theorem, 49
Fluid motion, 204
Force of attraction, 172, 202, 369, 372
Fourier coefficients, 380, 389
Fourier integral, 409, 413
  equation, 413
Fourier series, 362–414
  complex form of, 403, 405
  differentiation of, 405
  expansion in, 390
  integration of, 405, 409
  uniform convergence of, 408

Fresnel's integrals, 365
Function, 22, 58, 321
  bounds of, 106
  composite, 45, 47, 439
  continuity of, 31, 35, 39, 59, 392
  differentiable, 43
  even, 391
  homogeneous, 75
  implicit, 68, 415–440
  integrable, 109, 110, 335*n.*
  jump in, 34
  multiple-valued, 22
  normal, 388
  odd, 391
  orthogonal, 388
  periodic, 379*n.*
  piece-wise continuous, 34
  sectionally continuous, 34
  single-valued, 22
  uniformly continuous, 38
Functional dependence, 423, 433
Functional determinant, 153, 157, 167, 200, 421, 438
Fundamental theorem of integral calculus, 110, 118, 120

G

Gamma functions, 372
Gauss's test for series, 230
Gauss's theorem, 169*n.*
Geometric series, 214
Gradient, 78
Gravitational potential (*see* Potential)
Gravity, center of, 137, 159, 165
Green's theorem, 167, 170, 172, 181

H

Helix, 85, 86
Homogeneous functions, 75
Hyperbolic functions, 306
Hyperbolic paraboloid, 325

I

Implicit functions, 68, 415–440
  differentiation of, 71, 416
  existence theorem on, 425
  higher derivatives of, 89
Improper integrals, 335–377
  multiple, 365
Indefinite integrals, 119
Indeterminate forms, 54
Induction, mathematical, 9, 431
Infinite integrals, 335, 347
  tests for convergence of, 341, 349, 367
    (*See also* Improper integrals)
Infinite series, 209–266
  absolute convergence of, 236, 240, 263, 268
  addition of, 212, 240, 281
  alternating, 234
  of arbitrary terms, 233
  conditional convergence of, 236
  convergence of, 209, 233, 236
  differentiation of, 261
  division of, 285
  double, 246
  expansion in, 298, 383
  of functions, 247, 275
  geometric, 214
  integration of, 258, 309
  multiplication of, 242, 281
  of orthogonal functions, 389
  of positive terms, 233
  of power functions, 267–334
    algebra, 280
    applications, 291–334
    calculations with, 285
    expansion in, 291
    integration, 309
    reversion, 289
    tests for convergence, 269
    uniqueness theorem on, 279
  remainder in, 234
  sum of, 209, 395
  tests for convergence, 215, 225, 237, 262, 269
  uniform convergence of, 247, 252, 275
    (*See also* Fourier series)
Infinitesimal, 64*n.*
Infinity, 16
Integrable function, 109, 335*n.*
Integral equation, 413

Integral test for series, 225
Integrals, applications of, 126, 199
  change of variables in, 124, 147, 155, 199
  containing parameters, 121, 353
  of Dirichlet, 387, 362
  definite, 99, 101
  differentiation of, 121, 353
  divergent, 335
  elliptic, 117, 129, 312, 313, 352
  evaluation of, 110, 118, 123, 131, 144, 178, 310, 357
  of Fourier, 409, 413
  improper, or infinite, 335, 365
    (*See also* Improper integrals)
  indefinite, 119
  integration under sign, 353
  line, 174
  multiple, 130, 139, 142
  Riemannian, 104
  surface, 161, 196
  tests for convergence, 341
  transformation of (*see* Green's theorem; Stokes's theorem)
  uniform convergence of, 355, 367
Integration, of Fourier series, 405
  under integral sign, 353
  by parts, 353
  region of, 131
  in series, 309
  of series, 258, 309
Interchange of order of differentiation, 87
Intervals, Abel's theorem on, 268
  of convergence, 248, 267, 269
  extension of, 401
  open and closed, 22
Inverse transformations, 421
Inversion transformation, 440
Iterated integrals, 136

J

Jacobians, 150, 153, 157, 167, 200, 421
  properties of, 438
Jump in function, 34

K

Kummer's test, 228

L

Lagrange's multipliers, 329, 331
Lagrangian form of remainder, 292
Laplace's equation, 208, 372
Law of the mean, 291
    (*See also* Mean-value theorem)
Leibnitz's theorem on series, 234
Lemniscate, 128
Length of curve, 127
L'Hospital's rule, 54
Liapunoff, A., 383
Limiting points, 17
Limits, 6, 16, 58
  existence of, 10
  of functions, 23
  iterated, 61
  left- and right-hand, 24
  repeated, 61
  theorems on, 26
  upper and lower, 16, 18
Line integrals, 174–208
  applications of, 171, 199
  properties of, 185, 198
  in space, 195
  transformation of, 181, 196
Logarithm, base of, 10, 28, 31
Logarithmic spiral, 128
Lower bound, 22, 23, 106

M

Maclaurin's formula, 295, 318, 320
Maclaurin's series, 290
Maxima and minima, 315, 321, 327, 332
Mean-value theorem, 51
    (*See also* Law of the mean)
Mean-value theorems for integrals, 113, 115, 131
Mertens, F., 244
Minima (*see* Maxima and minima)
Minimax, 326
Moment of inertia, 138, 145, 151, 160, 166

Monotonicity, 14, 15
Multiple integrals, 130, 139, 142
  improper, 365
Multiplication of series, 244
Multiply connected regions, 184, 192

N

Neighborhood, 8
Newton's law of attraction, 172, 202
Newtonian potential, 172, 203, 370
Normal derivative, 78, 86, 87
Normal line, 80
Normal orthogonal functions, 388
Number system, 2
Numbers, 1
  complex, 3
  rational, 1
  real, 2

O

Odd function, 391
Order of, 406
Ordering, 1
Orthogonal functions, 388
Ostrogradsky's theorem, 169*n.*

P

Parameter, integrals containing, 121
Partial derivatives, 62
Partial sums, 209, 248
Partition, 2
Periodic function, 379*n.*
Plane, tangent, 80
Points of condensation, 17
Polar coordinates, 148, 152, 157
Potential, 172, 203, 206, 370
Power series, 267–290
  applications of, 291–334
Primitive, 119
Principal part of increment, 64
Probability integral, 357

Q

Quotient of power series, 282

R

Raabe's test, 230*n.*, 233, 303*n.*
Radius of convergence, 269
Ratio test, 217, 220, 237
Rearrangement of series, 240
Regions, simply and multiply connected, 184, 192, 196
Remainder, 234
  in Taylor's series, 292, 293
Riemann integral, 99
  upper and lower, 108
Riemann zeta-function, 364, 365
Rolle's theorem, 50
Root test, 216, 220, 237

S

Schwarz's theorem, 87
Sequences, 3
  bounded, 5
  convergent, 6, 14
  divergent, 16
  limiting points of, 17
  limits of, 6
  monotone, 14
  null, 5
Series (*see* Fourier series; Infinite series)
Simply connected regions, 184, 196
Simultaneous equations, existence theorem for, 430
Sine and cosine series, 397
Solid angle, 171
Space curve, 83
  equation of tangent to, 85
  length of, 86
Spherical coordinates, 157
Spiral, logarithmic, 128
Stokes's theorem, 196
Stream function, 206
Summation of series, 395
Surface integrals, 161, 196
Surface of revolution, 128

T

Tangent line, 70
Tangent plane, 80

Taylor's formula, 292, 293, 317
  applications of, 298, 322
Taylor's series, 296
Tests, for integrals, 341, 349, 367
  for series, 215, 225, 237, 262, 264
Total differential, 64
  (*See also* Exact differential)
Transformation, of coordinates, 153
  of integrals, 167, 170, 172, 181, 196
  of inversion, 440
Trigonometric series (*see* Fourier series)
Triple integrals, 142
  improper, 367

U

Undetermined multipliers, 328
Uniform continuity, 37, 60
Uniform convergence, of integrals, 355, 367
  of series, 247, 262, 264, 408
Upper bound, 23
  of a function, 106

V

Velocity potential, 206
Vicinity, 8
Volume, element of, 156, 158, 159
  of revolution, 128, 143, 155

W

Wallis's formula, 311
Weierstrass test, for integrals, 355
  for series, 262
Work, as a line integral, 201

Y

Young's theorem, 89*n*.

Z

Zeta-function, 364, 365